

UNIVERSITÀ DEGLI STUDI DEL PIEMONTE  
ORIENTALE “AMEDEO AVOGADRO”



UNIVERSITÀ DEL PIEMONTE ORIENTALE

Dipartimento di Scienze e Innovazione Tecnologica

Laurea Magistrale

*Corso di Laurea in Biologia indirizzo Biomedico-Biomolecolare*

Tesi di laurea

***“Analisi della predisposizione genetica e del microbiota intestinale  
mediante Next Generation Sequencing e Shotgun Sequencing in  
pazienti con polipi del colon-retto”***

Docente Relatore:

*Prof.ssa Irma Dianzani*

A handwritten signature in black ink, appearing to read 'Irma Dianzani'.

Candidato:

*Sara Cagliano*

N° matricola: 20029319

A handwritten signature in black ink, appearing to read 'Sara Cagliano'.

*Anno Accademico 2022/2023*

# Indice

<b>1. Contesto Generale</b> .....	4
<b>2. Titolo e riassunto</b> .....	6
<b>3. Introduzione</b> .....	9
3.1 Il Cancro del Colon-retto (CRC) .....	10
3.2 Patogenesi del CRC.....	11
3.3 Epidemiologia del CRC.....	13
3.4 Importanza dello stile di vita e della dieta nell'insorgenza del CRC.....	15
3.5 Predisposizione Genetica al CRC.....	17
3.5.1 <i>Sindrome di Lynch</i> .....	17
3.5.2 <i>Poliposi Adenomatosa familiare</i> .....	17
3.5.3 <i>Poliposi MUTYH-Associata</i> .....	18
3.5.4 <i>Poliposi amartomatose</i> .....	18
3.6 Microbiota Intestinale.....	19
<b>4. Scopo della tesi</b> .....	22
<b>5. Materiali e Metodi</b> .....	24
5.1 Reclutamento pazienti e raccolta campioni biologici.....	25
5.2 Estrazione del DNA.....	25
5.3 Quantificazione del DNA.....	26
5.4 Sequenziamento del DNA germinale mediante Next Generation Sequencing (NGS).....	27
5.5 Analisi <i>in silico</i> .....	34
5.6 Disegno Primer.....	34
5.7 PCR.....	35
5.8 Elettroforesi.....	36
5.9 Purificazione dell'amplificato.....	37
5.10 Preparazione delle reazioni di sequenza.....	38
5.11 Purificazione delle reazioni di sequenza.....	39
5.12 Sequenziamento di Sanger.....	39
5.13 Analisi delle sequenze.....	39
5.14 Conferma delle varianti sul DNA tumorale.....	39
5.15 Classificazione pazienti.....	40
5.16 Analisi del microbioma adeso al polipo e luminale (Shotgun Sequencing).....	40

<b>6. Risultati</b> .....	42
6.1 Risultati del sequenziamento del DNA germinale mediante NGS.....	43
6.2 Conferme delle varianti con sequenziamento di Sanger.....	46
6.2.1 DNA germinale e tumorale: paziente s129.....	46
6.2.2 DNA germinale e tumorale: paziente s219.....	47
6.2.3 DNA germinale e tumorale: paziente s242.....	48
6.2.4 DNA germinale e tumorale: paziente s243.....	48
6.2.5 DNA germinale e tumorale: paziente s293.....	49
6.2.6 DNA germinale e tumorale: paziente s294.....	49
6.3 Varianti patogeniche evidenziate nell'intera casistica dei pazienti.....	50
6.4 Classificazione dei pazienti e analisi del microbiota.....	52
<b>7. Discussione dei risultati</b> .....	55
<b>8. Bibliografia</b> .....	60

## **1. Contesto generale**

## **Attività di laboratorio**

Ricerca di base universitaria eseguita presso l'Università del Piemonte Orientale - Laboratorio di Patologia Genetica, Dipartimento di Scienze della Salute, diretto dalla Professoressa Irma Dianzani.

## **2. Titolo e Riassunto**

***“Analisi della predisposizione genetica e del microbiota intestinale mediante Next-Generation Sequencing e Shotgun Sequencing in pazienti con polipi del colon-retto”***

Il mio tirocinio di tesi si è svolto presso il laboratorio di Patologia Genetica del Dipartimento di Scienze della Salute dell'Università del Piemonte Orientale. Si è focalizzato sull'identificazione di varianti germinali patogeniche predisponenti all'insorgenza del cancro in pazienti affetti da tumori del colon-retto e sul confronto del microbiota intestinale nei pazienti con e senza predisposizione genetica.

Il cancro al colon-retto (CRC) è una delle patologie neoplastiche più diffuse a livello globale, sia nella popolazione maschile che femminile. Si tratta di un tumore ad origine multifatoriale, caratterizzato da fattori di rischio modificabili, tra cui dieta, obesità, mancanza di attività fisica, uso di tabacco ed uso moderato o pesante di alcol e fattori non modificabili, quali sesso, età e predisposizione genetica. Il 30% dei pazienti presenta familiarità, mentre solo l'8-10% presenta una sindrome ereditaria ad alto rischio di sviluppo neoplastico. In particolare, diversi geni possono essere responsabili, se mutati, dell'insorgenza del CRC, tra cui *APC* e i geni del mismatch repair *MLH1*, *MSH2*, *MSH6* e *PMS2*. Sempre più dati suggeriscono che il microbiota intestinale abbia un ruolo importante nello sviluppo del CRC, inducendo infiammazione e producendo metaboliti e sostanze genotossiche.

Durante il mio tirocinio mi sono occupata dell'analisi del genoma germinale di 63 pazienti con polipi intestinali utilizzando un pannello Next-Generation Sequencing custom che include 107 geni associati allo sviluppo di sindromi ad alta incidenza di cancro. Dall'analisi effettuata sono emerse 25653 varianti, di cui 1542 avevano frequenza allelica <0.01. Tra di esse sono state identificate le seguenti varianti patogeniche: c.1103G>A p.Gly368Asp nel gene *MUTYH*, c.1111C>T p.Arg371Trp nel gene *FANCF*, c.2570A>C p.Asp857Ala nel gene *MSH6* e c.244C>T p.Gln82\* nel gene *NTHL1*.

Sono state validate tramite sequenziamento di Sanger sul DNA germinale e dove disponibile sul DNA tumorale, per indagare l'eventuale perdita del secondo allele causata da mutazioni somatiche.

Inoltre, è stato effettuato lo studio del microbioma adeso al tumore, grazie ad una nuova metodica messa a punto dal gruppo di ricerca in cui ho svolto il tirocinio e basata sull'utilizzo di un tampone e-NAT® strofinato sul polipo rimosso durante la colonscopia.

Il sequenziamento metagenomico Shotgun ha permesso di identificare la composizione del profilo tassonomico, con la stima degli arricchimenti relativi a livello di specie, nel gruppo di

pazienti con predisposizione genetica rispetto ai pazienti non mutati. Tramite il sequenziamento del microbiota adeso alla mucosa del polipo (MAM) sono state identificate differenti firme batteriche nei pazienti sporadici e mutati; in particolare, nei pazienti mutati è stato riscontrato un arricchimento di *Fusobacterium nucleatum*, una specie batterica associata al CRC. Questi dati supportano l'ipotesi che pazienti con differenti background genetici abbiano diversi fattori di rischio legati al microbiota intestinale.

### **3. Introduzione**

### 3.1 Il Cancro del Colon-retto (CRC)

Il carcinoma del colon-retto è il terzo tumore più comune negli uomini e il secondo nelle donne, con quasi 2 milioni di nuovi casi nel 2020, e la seconda causa di decessi correlati al cancro in tutto il mondo con più di 700000 decessi all'anno (Sung *et al.*, 2021; Morgan *et al.*, 2023).

Allo sviluppo del CRC sono associati diversi fattori di rischio modificabili, tra cui la dieta, l'obesità, la mancanza di attività fisica, l'uso di tabacco e l'uso moderato o pesante di alcol e fattori non modificabili, quali sesso, età, predisposizione genetica e condizioni ereditarie come la sindrome di Lynch (Karen Simon, 2016; Thanikachalam *et al.*, 2019). Le malattie infiammatorie del tratto intestinale possono favorire l'insorgenza del CRC. Con Inflammatory Bowel Diseases (IBD) si fa riferimento ad una condizione infiammatoria cronica, che colpisce il tratto gastrointestinale e può evolvere in CRC (Majumder *et al.*, 2022). Una delle cause dell'IBD è la variazione della risposta immunitaria della mucosa intestinale ai batteri enterici (Shanahan *et al.*, 2001).

Il CRC spesso cresce lentamente e di solito non produce sintomi fino a quando non raggiunge una dimensione di diversi centimetri, bloccando il passaggio delle feci e portando a crampi, dolore o sanguinamento. La maggior parte dei tumori del colon si sviluppa attraverso un processo multistep, chiamato "sequenza adenoma-carcinoma", che prevede una serie di alterazioni istologiche, morfologiche e genetiche che si accumulano nel tempo (Karen Simon, 2016).

In genere, il CRC ha origine a partire da un polipo, un tumore benigno che si sviluppa dallo strato mucoso del colon o del retto (Alteri *et al.*, 2022). Quando le cellule epiteliali intestinali iper-proliferative perdono la loro organizzazione e struttura, acquistano la capacità di formare gli adenomi, che possono quindi diventare carcinomi, invadere la sottomucosa e generare metastasi a distanza (Lucas *et al.*, 2017). Solo una piccola parte dei polipi acquisisce caratteristiche maligne e la progressione completa da polipo a carcinoma richiede diversi anni o anche più di un decennio (Karen Simon, 2016).

I CRC possono originare dai polipi adenomatosi, che sono i precursori più comuni del cancro, o serrati, cosiddetti a causa del loro aspetto dentellato al microscopio. I primi sono classificati in tubulari, tubulovillosi e villosi, mentre i secondi sono ulteriormente suddivisi in polipi serrati sessili (SSP), adenomi serrati tradizionali (TSA) e polipi iperplastici (HPP) (Alteri *et al.*, 2022).

Il rischio che gli adenomi progrediscono in CRC aumenta con l'aumentare della grandezza del polipo, in particolare gli adenomi che presentano dimensioni  $\geq 1$  cm e componente villosa  $\geq$

25% o mostrano displasia di alto grado rappresentano la via di progressione del 60%-70% dei CRC (Karen Simon, 2016).

In rapporto alla localizzazione anatomica, i tumori del colon sono suddivisi in:

- prossimali (o del lato destro), quando hanno origine nel cieco, nel colon ascendente o nel trasverso,
- distali (o del lato sinistro), se sorgono nel colon discendente o nel colon sigmoideo,
- del retto (Testa *et al.*, 2018).

### 3.2 Patogenesi del CRC

La progressione neoplastica è il risultato dell'accumulo di alterazioni genetiche ed epigenetiche, che trasformano le cellule epiteliali ghiandolari normali generando prima polipi e poi adenocarcinomi invasivi (Grady *et al.*, 2014). La sommatoria di questi eventi fa sì che l'epitelio della mucosa intestinale acquisisca le caratteristiche indispensabili per la crescita incontrollata: l'accumulo di mutazioni in geni specifici e la conseguente deregolamentazione della segnalazione cellulare determinano lo sviluppo di comportamenti caratteristici del cancro quali iper-proliferazione cellulare, ridotta differenziazione, immortalizzazione, angiogenesi, invasione e metastasi (Grady *et al.*, 2014). Sono stati proposti due modelli di evoluzione del tessuto normale in tumorale: il pathway classico e il pathway alternativo, caratterizzati da diversi meccanismi di instabilità genomica, quali instabilità cromosomica (CIN), instabilità dei microsatelliti (MSI) e CpG island methylator phenotype (CIMP) (Mármol *et al.*, 2017).

Il pathway classico vede lo sviluppo di polipi adenomatosi tubulari o tubulovillosi che evolvono in adenocarcinomi in seguito all'acquisizione di mutazioni (Grady *et al.*, 2014) a livello delle cellule presenti sul fondo della cripta intestinale (Testa *et al.*, 2018). La via di instabilità cromosomica (CIN), generalmente associata agli adenomi tradizionali, è osservata nell'85%-90% di tutti i tumori sporadici (Grady *et al.*, 2008; Keum *et al.*, 2019). Essa è caratterizzata dalla presenza di anomalie cromosomiche numeriche o strutturali e frequente perdita di eterozigosi (LOH) nei loci dei geni oncosoppressori. Il pathway di CIN è associato alla perdita o all'inattivazione di *Adenomatous polyposis coli* (APC), che provoca l'attivazione costitutiva della via di segnalazione di Wnt/ $\beta$ -catenina (Bian *et al.*, 2020).

Il pathway di Wnt è responsabile del mantenimento, differenziamento e della proliferazione delle cellule staminali enteriche (Barker *et al.*, 2014). In assenza di un ligando Wnt, la  $\beta$ -catenina citoplasmatica è fosforilata dal complesso di distruzione (costituito da APC, dalla

proteina di inibizione dell'asse (AXIN), dalla glicogeno sintasi chinasi 3 (GSK3) e dalla casein chinasi 1 (CK1)), che induce la continua eliminazione di  $\beta$ -catenina, così impedendo il suo trasporto nel nucleo. In questo modo l'espressione dei geni target di Wnt è spenta (Bian *et al.*, 2020). La presenza di un ligando Wnt, invece, blocca o interrompe la formazione del complesso di Axin/GSK3/APC, inibendo così la degradazione della  $\beta$ -catenina e permettendone l'accumulo a livello citoplasmatico. La  $\beta$ -catenina citoplasmatica viene quindi trasferita nel nucleo, dove permette l'espressione dei geni bersaglio Wnt (Zhao *et al.*, 2022).

Il pathway alternativo, osservato in circa il 10-15% dei pazienti con CRC sporadico (Keum *et al.*, 2019), comporta la formazione di polipi iperplastici che possono evolvere in serrati o sessili serrati ed infine in tumori maligni (Grady *et al.*, 2014). Questo pathway è caratterizzato dalle mutazioni dei geni *BRAF* e *KRAS*, e in seguito dal CIMP (Puccini *et al.*, 2017; Testa *et al.*, 2018). Circa il 30-45% dei pazienti con CRC presentano la mutazione del gene *KRAS*, mentre il 5-20% dei pazienti portano la mutazione del gene *BRAF* (Uchida *et al.*, 2022). *KRAS* e *BRAF* codificano per proteine implicate nella regolazione della sopravvivenza cellulare; di conseguenza, le mutazioni gain-of-function di *KRAS* e di *BRAF* tengono attive queste vie (Uchida *et al.*, 2022).

Il tumore del colon causato da CIMP è caratterizzato da un'ipermetilazione di tutto il genoma in specifiche regioni del DNA, chiamate isole CpG. Si tratta di sequenze di circa 200-500 basi di lunghezza con un contenuto di citosina e guanina superiore al 50%. Di solito, le isole CpG si sovrappongono alla regione promotore in 60-70% dei geni e tendono ad essere protette dalla metilazione (Puccini *et al.*, 2017). La loro metilazione aberrante causa il silenziamento trascrizionale e può portare all'inattivazione di oncosoppressori (Toyota *et al.*, 1999; Puccini *et al.*, 2017).

I CRC del pathway alterativo possono presentare MSI. I tumori con mutazioni in *BRAF*, CIMP e stabilità dei microsatelliti (MSS) sono associati a prognosi peggiore rispetto a quelli con alta MSI (Testa *et al.*, 2018).

MSI è un fenotipo ipermutabile determinato dalla perdita di attività del sistema del riparo del DNA del mismatch repair (MMR) a causa di mutazioni o di silenziamento epigenetico dei geni *MLH1*, *MSH2*, *MSH6* e *PMS2*, che rendono le regioni dei microsatelliti suscettibili a mutazioni del DNA (Puccini *et al.*, 2017). I due meccanismi CIMP e MSI sono correlati poiché l'ipermetilazione delle isole CpG è uno dei meccanismi che inattiva i geni del MMR (Keum *et al.*, 2019).

### 3.3 Epidemiologia del CRC

Il cancro costituisce una delle principali cause di decesso e rappresenta una importante barriera all'incremento dell'aspettativa di vita in tutto il mondo. Nel 2019, secondo le valutazioni dell'Organizzazione Mondiale della Sanità (OMS), il cancro si posiziona come la prima o seconda causa di morte prima dei 70 anni in 112 dei 183 paesi presi in considerazione, mentre occupa la terza o quarta posizione in altri 23 paesi (Sung *et al.*, 2021).

In particolare, il CRC è il terzo cancro diagnosticato e la seconda causa principale delle morti per cancro a livello mondiale (Morgan *et al.*, 2023; Sung *et al.*, 2021). Nel 2020 sono stati stimati circa 2 milioni di nuovi casi di CRC e circa 1 milione di decessi da esso provocati, corrispondenti al 10,7% e 9,5% di tutti i nuovi casi di cancro e decessi nel mondo (Morgan *et al.*, 2023). Dal 1994 l'incidenza del CRC negli individui di età inferiore ai 50 anni è aumentata del 2% all'anno, mentre l'incidenza generale di CRC in Stati Uniti e Europa sta diminuendo grazie alla diagnosi precoce (Mauri *et al.*, 2018).

L'incidenza del CRC varia in base al sesso: l'incidenza nel sesso femminile è del 9,4%, con mortalità del 9,5%, rappresentando il secondo tumore più frequente nella popolazione femminile, dopo il cancro al seno; invece, l'incidenza stimata nel sesso maschile è del 10,6%, con mortalità del 9,3%, rappresentando il terzo tumore per incidenza nella popolazione maschile (Sung *et al.*, 2021, GLOBOCAN 2020) (Fig.1).

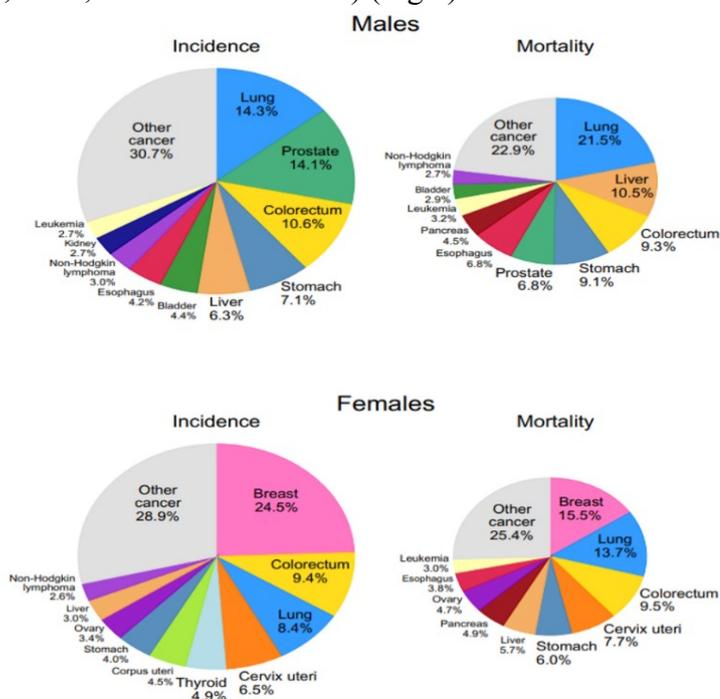


Figura 1. Incidenza e mortalità per CRC in maschi e femmine (GLOBOCAN 2020)

Anche l'età influenza l'incidenza del CRC. In particolare, l'incidenza di CRC sale nella popolazione con età superiore ai 50 anni (Araghi *et al.*, 2019), con maggior rilevanza per il CRC sporadico che presenta un picco tra i 70 e i 75 anni di età (Yamagishi *et al.*, 2016). A livello globale molti paesi negli ultimi anni hanno riscontrato un aumento dell'incidenza dello sviluppo del CRC in una popolazione più giovane ed una diminuzione o mantenimento dell'incidenza nella popolazione over 50, come in Germania, Stati Uniti, Svezia, Canada, Regno Unito ed India. Solo l'Italia ha registrato l'andamento inverso di questa tendenza (Wong *et al.*, 2021). A differenza del CRC sporadico, le forme di poliposi familiare si sviluppano in età precoce, con diagnosi di CRC a partire dai 40 anni (Lynch *et al.*, 2003).

L'incidenza e la mortalità del CRC presentano variazioni importanti anche a livello geografico. Dai dati riportati sul Globocan Cancer Observatory per il 2020 si può osservare che l'Asia presenta le percentuali più alte, rispettivamente 52,3% e 54,2%, sia per l'incidenza che per la mortalità in entrambi i sessi; seguita poi dall'Europa (26,9% per incidenza e 26,2% per mortalità), dal Nord America per l'incidenza (9,3%) e dall'America Latina e dai Caraibi per mortalità (7,4%) (GLOBOCAN 2020) (Fig. 2). Sebbene i paesi con basso indice di sviluppo, come l'Africa e Asia Meridionale, presentino le percentuali d'incidenza e mortalità più basse, è stato registrato un incremento di casi negli ultimi anni, che riflette i cambiamenti legati allo stile di vita e alla dieta (Araghi *et al.*, 2019). Lo sviluppo dello stato socioeconomico fa sì che aumenti il consumo di alimenti di origine animale e che si instauri uno stile di vita sedentario, con la diminuzione dell'attività fisica e aumento del peso corporeo. (Thanikachalam *et al.*, 2019; Sung *et al.*, 2021).

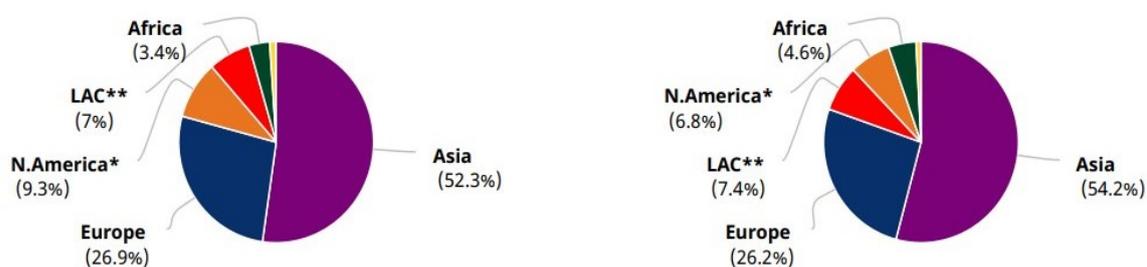


Figura 2. Incidenza e la mortalità del CRC per zone geografiche (GLOBOCAN 2020). Abbreviazioni= \*N. America: Nord America, \*\*LAC: America Latina e Caraibi

Un metodo efficace per ridurre l'incidenza del CRC è stato lo sviluppo di programmi di screening, oltre all'incremento dell'attività fisica e della dieta, (Ladabaum *et al.*, 2020). In molti paesi, infatti, l'implementazione di interventi preventivi e curativi è riuscita a mantenere stabile l'incidenza del CRC negli ultimi anni (Sung *et al.*, 2021; Siege *et al.*, 2023). Tuttavia, i dati sulle prospettive future annunciano un incremento a 3,2 milioni di casi e 1,6 milioni di morti entro il 2040, con un aumento rispettivamente del 63% e del 73% rispetto ai valori attuali (Morgan *et al.*, 2023; Xi *et al.*, 2021).

### **3.4 Importanza dello stile di vita e della dieta nell'insorgenza del CRC**

L'associazione tra attività fisica e rischio di sviluppo del cancro del colon è supportata da molti studi epidemiologici osservazionali, valutati e riassunti dal Physical Activity Guidelines for Americans (PAGA) 2018 Report e dal World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR), come linee guida al fine di ridurre il rischio (Friedenreich *et al.*, 2020; Wolin *et al.*, 2009). Già a partire dai primi anni '70 era stata ipotizzata l'associazione tra stile di vita e cancro al colon. Con l'adozione di uno stile di vita sempre più occidentale, negli Stati Uniti si è osservato un incremento di malattie non infettive dell'intestino, come i polipi adenomatosi, nelle nuove generazioni di afroamericani rispetto alle generazioni precedenti (Burkitt *et al.*, 1971). I dati dimostrano una riduzione del 24% del rischio di sviluppare CRC in soggetti con stile di vita attivo rispetto a soggetti con stile di vita sedentario, che comporta aumento di peso e obesità (Wolin *et al.*, 2009; Song *et al.*, 2020). A sostegno di questa ipotesi, i pazienti che hanno subito una significativa perdita di peso e chirurgia bariatrica presentano una riduzione del rischio di CRC nel periodo di follow-up di 5-12,5 anni rispetto a prima della perdita di peso (Friedenreich *et al.*, 2020).

Anche la dieta e il consumo di diversi alimenti possono essere correlati ad una variazione del rischio di neoplasia coloretale (Song *et al.*, 2015): fibre, frutta, verdura e grano sono associati ad un minor rischio di CRC, mentre carni rosse e processate sono legate ad un rischio maggiore (Pan *et al.*, 2018; Song *et al.*, 2015). Le fibre hanno un ruolo protettivo a livello intestinale, in quanto riducono il tempo di transito fecale, partecipano al mantenimento delle cellule epiteliali enteriche e vengono fermentate dai batteri anaerobi, creando acidi grassi a catena corta, quali acetato, propionato e butirato che inibiscono la progressione neoplastica (Burkitt *et al.*, 1971; Pan *et al.*, 2018; Song *et al.*, 2015). La frutta e la verdura possono avere un ruolo protettivo contro il CRC: contengono alti livelli di molte sostanze anticancerogene, in particolare composti polifenolici, flavonoidi, vitamine del gruppo B e minerali (Pan *et al.*, 2018; Song *et al.*, 2015). Al contrario, il consumo di carne rossa e processata è associato all'aumento del

rischio di CRC (Abid *et al.*, 2014). Si osserva un aumento del 20% del rischio di sviluppo del CRC correlato all'aumento di 100g/die di carne rossa e 50g/die di carne processata, e si ipotizza che il rischio aumenti in modo proporzionale all'assunzione di questi alimenti (Song *et al.*, 2015). La metabolizzazione da parte dei batteri enterici delle carni rosse provoca la produzione di metaboliti coinvolti nel CRC, come ad esempio la colina e la carnitina. Esse sono ampiamente contenute nelle carni rosse e sono precursori della formazione di TMA (trimetilammina), trasformata poi in TMAO (trimetilammina-N-ossido), che induce infiammazione e proliferazione cellulare (Song *et al.*, 2015; Song *et al.*, 2020; Yang *et al.*, 2019). Altri derivati della carne rossa e processata (o del loro processo di cottura) che aumentano il rischio di CRC sono: le ammine eterocicliche (HCAs), gli idrocarburi aromatici policiclici (PAHs), composti N-nitrosi, grassi. Inoltre, l'altro contenuto di grassi conduce ad un'aumentata secrezione di acidi biliari, che vengono convertiti in acidi biliari secondari cancerogeni dal microbiota. (Song *et al.*, 2020; Baena *et al.*, 2015).

Anche il fumo di tabacco e l'eccessivo consumo di alcol sono considerati fattori di rischio per la neoplasia coloretale (Bailie *et al.*, 2017). È stato osservato un aumento del 60% del rischio di CRC nei fumatori, con associazione alle forme tumorali che presentano MSI (Brenner *et al.*, 2018). Si stima che il consumo di alcool incrementi il rischio d'insorgenza di CRC del 20% se moderato, ma fino al 50% se elevato (Kovacs *et al.*, 2009; Fedirko *et al.*, 2011).

### **3.5 Predisposizione Genetica al CRC**

Il 30% dei pazienti presenta familiarità per il CRC e una storia familiare positiva è un fattore di rischio per il CRC (Stoffel *et al.*, 2014; Olkinuora *et al.*, 2021). Il rischio associato alla familiarità è proporzionale al numero di membri con la stessa diagnosi, alla loro età al momento della diagnosi e al grado di parentela (Hryhorowicz *et al.*, 2022; Tian *et al.*, 2020). Il 2-8% dei casi di CRC è causato da varianti patogene germline in geni associati a sindromi ad alto rischio del cancro (Valle *et al.*, 2019), quali la sindrome di Lynch (LS), la poliposi adenomatosa familiare (FAP), la poliposi MUTYH-associata (MAP) e le poliposi amartomatose.

#### **3.5.1 Sindrome di Lynch**

La LS è la più comune sindrome ereditaria autosomica dominante associata a predisposizione a neoplasie, tra cui la più comune è il CRC (Stoffel *et al.*, 2014; Hryhorowicz *et al.*, 2022). Circa il 3% delle diagnosi complessive di CRC è associato alla LS (Lynch *et al.*, 2009). Nei pazienti con LS la cancerogenesi risulta accelerata rispetto al cancro sporadico, da 1 a 3 anni circa, con un'età d'esordio precoce (età media di 45 anni) (Lynch *et al.*, 2009; Hryhorowicz *et al.*, 2022). La LS è causata da una mutazione in uno dei quattro geni del mismatch repair (MMR) *MLH1*, *MSH2*, *MSH6*, *PMS2* o del gene a monte di *MSH2*, ovvero *EPCAM* (Ligtenberg *et al.*, 2009; Kovacs *et al.*, 2009; Stoffel *et al.*, 2014). Delezioni del gene *EPCAM*, che non è coinvolto nel MMR, portano al silenziamento epigenetico di *MSH2* mediante ipermetilazione (GeneReviews). Le mutazioni germline più frequenti ricadono nei geni *MLH1* (42%) e *MSH2* (33%), seguiti da *MSH6* e *PMS2* (18-7,5%) (Plazzer *et al.*, 2013; GeneReviews). Le mutazioni che provocano l'inattivazione dei geni MMR sono mutazioni frameshift, nonsense, missenso (le più comuni per *MLH1* e *MSH2*), mutazioni sui siti di splicing, delezioni o riarrangiamenti (Lung *et al.*, 2015).

#### **3.5.2 Poliposi Adenomatosa familiare**

Il 90% dei casi di FAP, disturbo ereditario autosomico dominante, è dovuto a mutazioni germinali del gene *APC*, situato sul cromosoma 5 (Lung *et al.*, 2015; Stoffel *et al.*, 2014; Aelvoet *et al.*, 2022). La FAP rappresenta circa l'1% dei casi totali di CRC ed è una patologia che porta allo sviluppo di centinaia o migliaia di polipi nel colon e/o nel retto già durante l'età adolescenziale (Lung *et al.*, 2015; Aelvoet *et al.*, 2022). La FAP è caratterizzata da una progressione adenoma-carcinoma molto simile al cancro al colon sporadico (15-20 anni).

Il rischio che evolva in cancro intorno ai 40 anni se il paziente non viene trattato è del 100% (Bisgaard *et al.*, 1994; Aelvoet *et al.*, 2022). La gravità della malattia varia in base al tipo di mutazione del gene *APC* colpiti dalla mutazione. La poliposi adenomatosa familiare attenuata (AFAP), è caratterizzata da un minor numero di polipi rispetto alla FAP (<100). FAP e AFAP sono causate da mutazioni patogeniche in eterozigosi in diversi domini del gene *APC* e con effetto differente (Lung *et al.*, 2015; Nielsen *et al.*, 2007).

### **3.5.3 Poliposi *MUTYH*-Associata**

La MAP è una sindrome responsabile di meno dell'1% di tutti i CRC ed è caratterizzata dallo sviluppo di più di 500 polipi (Aelvoet *et al.*, 2022; Casper *et al.*, 2018). L' ereditarietà è autosomica recessiva (Testa *et al.*, 2018). *MUTYH* codifica per una DNA glicosilasi, coinvolta nella riparazione per escissione di basi (BER), responsabile dell'eliminazione delle adenine appaiate in modo scorretto (Testa *et al.*, 2018; Brinkmeyer *et al.*, 2015). Le mutazioni bialleliche con perdita di funzione nel gene *MUTYH* determinano l'accumulo di trasversioni somatiche G > T in molti geni, tra cui *APC* e *K-RAS* (Aelvoet *et al.*, 2022; Brinkmeyer *et al.*, 2015).

### **3.5.4 Poliposi amartomatose**

Le poliposi amartomatose, sebbene rappresentino meno dello 0,5% di tutti i casi di CRC, sono associate allo sviluppo di diversi tumori, tra cui il CRC (Stoffel *et al.*, 2014; Samadder *et al.*, 2015). La formazione di polipi amartomatosi si osserva in diverse sindromi, tra cui sindrome di Peutz-Jeghers (PJS), la poliposi giovanile (JPS), la malattia di Cowden e la sindrome di Bannayan– Ruvalcaba–Riley (Lynch *et al.*, 2003). In particolare, la PJS e al JPS possono condurre alla formazione di CRC.

La PJS porta alla formazione di più polipi amartomatosi nell'intestino, causa di dolore addominale, sanguinamento gastrointestinale, anemia, ostruzione intestinale e intussuscezione (Lung *et al.*, 2015; Stoffel *et al.*, 2014). Questa sindrome è provocata, in circa il 50-70% dei pazienti, da una mutazione del gene oncosoppressore *STK-11*, che partecipa nella via di mTor e che codifica per la serina-treonina chinasi 11 (Stoffel *et al.*, 2014; Hearle *et al.*, 2006; Lynch *et al.*, 2003).

La JPS è una condizione autosomica dominante, caratterizzata da polipi giovanili multipli, più comuni nello stomaco e nel colon, che aumentano del 40-50% il rischio di tumori maligni gastroenterici (Lung *et al.*, 2015; Stoffel *et al.*, 2014). Nel 40-60% dei casi la JPS è caratterizzata

da mutazioni dei geni *SMAD4* o *BMPRIA*, che codificano per proteine coinvolte nel pathway di TGF- $\beta$  (Stoffel *et al.*, 2014; Calva-Cerqueira *et al.*, 2010).

### 3.6 Microbiota Intestinale

Il microbiota intestinale, l'insieme di tutti gli organismi che popolano l'intestino, è composto da circa 2000 specie di batteri, funghi, virus, Archaea e protozoi (Lucas *et al.*, 2017). In condizioni di salute il microbiota intestinale crea un rapporto simbiotico con l'ospite, e per questo risulta avere un ruolo primario in molte funzioni, nell'utilizzo degli alimenti come fonte di energia, nel trofismo dell'epitelio intestinale, nella protezione dagli agenti patogeni e nella risposta immune (Lucas *et al.*, 2017; Song *et al.*, 2020).

Un disequilibrio del rapporto microbiota-ospite può condurre ad uno stato di disbiosi, ovvero ad un'alterazione del microbiota intestinale, che a sua volta può portare a varie malattie, tra cui il CRC (Cheng *et al.*, 2020; Song *et al.*, 2020). Molti studi hanno evidenziato come la composizione del microbiota intestinale sia modificata in pazienti con CRC rispetto ai soggetti sani. In generale, nei pazienti si osserva una minore diversità e ricchezza batterica (Cheng *et al.*, 2020). Inoltre, il microbiota intestinale dei pazienti risulta alterato anche in presenza di adenomi. Quindi, le diverse fasi di cancerogenesi risultano correlate a specifiche variazioni del microbiota, che possono essere utilizzate come biomarcatori per lo screening di CRC (Cheng *et al.*, 2020; Song *et al.*, 2020).

I phyla batterici più rappresentati nel microbiota intestinale normale sono Firmicutes, Bacteroidetes, Actinobacteria, e Proteobacteria. Alcune specie batteriche risultano associate all'insorgenza e alla progressione del CRC (Cheng *et al.*, 2020; Castellarin *et al.*, 2012; Kostic *et al.*, 2013; Marchesi *et al.*, 2011; Srikanth *et al.*, 2008; Wei *et al.*, 2010).

Tra le specie associate al CRC ci sono *Fusobacterium nucleatum*, *Bacteroides fragilis*, *Escherichia coli* (Flanagan *et al.*, 2014; Boleij *et al.*, 2015; Bonnet *et al.*, 2014). *Fusobacterium nucleatum*, batterio Gram negativo anaerobio e patogeno opportunisto, è arricchito nei pazienti con CRC in stadio avanzato o tempi di sopravvivenza ridotti (Song *et al.*, 2020; Flanagan *et al.*, 2014). Il suo potenziale cancerogeno è associato all'attivazione di due fattori di virulenza: FadA, fattore di adesione che permette al batterio di invadere il tessuto epiteliale dell'ospite più facilmente, e Fap2, proteina di adesione galattosio-sensibile che favorisce l'emo-agglutinazione e l'adesione cellulare (Liu *et al.*, 2019; Han *et al.*, 2000; Han *et al.*, 2004; Copenhagen-Glazer *et al.*, 2015). Attraverso l'adesione di FadA, il batterio favorisce anche il reclutamento delle cellule immunitarie tumorali e la formazione di un microambiente che promuove la progressione neoplastica, caratterizzato da citochine pro-infiammatorie, quali

IL-6, IL-8, IL-10, TNF- $\alpha$  (Flanagan *et al.*, 2014; Rubinstein *et al.*, 2013). Inoltre, questo fattore partecipa all'attivazione del pathway di segnalazione della  $\beta$ -catenina, grazie all'interazione con la E-caderina, sostenendo la cancerogenesi del CRC (Rubinstein *et al.*, 2013). Una maggior abbondanza di questo batterio nei pazienti con CRC in stadi avanzati può essere ricondotta alla presenza della proteina Fap2, che può legare l'acetilgalattosamina, polisaccaride over-espresso nel tumore del colon-retto (Abed *et al.*, 2016). Inoltre, Fap2 interagisce con TIGIT (T-cell immunoglobulin and ITIM domain), recettore inibitorio espresso sulla superficie dei linfociti T e delle cellule NK (natural killer), promuove la riduzione della tossicità delle cellule NK e la soppressione della risposta immunitaria contro il tumore (Gur *et al.*, 2015). Il batterio anaerobio *Bacteroides fragilis* rappresenta meno dell'1% del microbiota intestinale ed è stato associato alla patogenesi del CRC per via della produzione di un'enterotossina, chiamata tossina metalloproteasi zinco dipendente (BFT) (Song *et al.*, 2020; Alhinai *et al.*, 2019; Huang *et al.*, 2011). La BFT, contribuisce allo sviluppo della risposta infiammatoria della mucosa intestinale, in quanto induce un aumento dell'espressione dell'interleuchina 17 (IL-17) (Tjalsma *et al.*, 2012). La tossina permette l'aumento della permeabilità intestinale e stimola l'attivazione del pathway Wnt/ $\beta$ -catenina induce il clivaggio dell'E-caderina. Inoltre, promuove l'espressione di STAT3 e IL-6, che sono fattori pro-proliferativi, anti-apoptotici e pro-angiogenici, fondamentali per la progressione tumorale (Housseau *et al.*, 2010).

In pazienti con malattie infiammatorie intestinali e con CRC è stato osservato un aumento dei livelli di *Escherichia coli* rispetto ai soggetti sani (Song *et al.*, 2020). In particolare, è arricchito un ceppo di *E.coli* contenente il cluster di geni per la polichetide sintasi, che regola la produzione di una genotossina, la colibactina (Song *et al.*, 2020). Questa tossina induce la rottura dei filamenti di DNA, favorendo l'insorgenza di aberrazioni cromosomiche (Cuevas-Ramos *et al.*, 2010; Alhinai *et al.*, 2019).

Il microbiota intestinale è implicato nella cancerogenesi sia direttamente che indirettamente, con la produzione di metaboliti e sostanze genotossiche da parte dei batteri (Scott *et al.*, 2019; Kosumi *et al.*, 2018; Louis *et al.*, 2014).

I batteri del microbiota, tramite la fermentazione delle fibre non digerite, producono acidi grassi a catena corta (SCFAs) come propionato, acetato e butirato (Zeng *et al.*, 2020; Kuwahara *et al.*, 2014); quest'ultimo ha effetti opposti sulla proliferazione delle cellule del colon sane e neoplastiche. Il butirato, agendo come fonte energetica, va a stimolare la proliferazione dei colonociti sani; al contrario, i colonociti neoplastici hanno difficoltà a metabolizzare il butirato per l'effetto Warburg. Il butirato, tende così ad accumularsi nelle cellule tumorali, agendo

come inibitore dell'istone deacetilasi, che permette un aumento dell'espressione di molecole pro-apoptotiche e una riduzione di quelle anti-apoptotiche (Donohoe *et al.*, 2014). Vari studi ipotizzano un effetto opposto sulla cancerogenesi in soggetti con diverso background genetico (Wang *et al.*, 2012; Wu *et al.*, 2013; Li *et al.*, 2016). In particolare, sono stati osservati effetti protettivi del butirrato nei topi con tumori indotti attraverso azossimetano e destrano sodio solfato (AOM/DSS) e, al contrario, effetti favorevoli alla cancerogenesi nei topi *Apc*<sup>Min/+</sup> *Msh2*<sup>-/-</sup>, portatori di mutazioni germline, che simulano le LS e la FAP, in cui il butirrato favorisce la proliferazione cellulare (Belcheva *et al.*, 2014; Donohoe *et al.*, 2014).

La disbiosi può essere contrastata tramite la somministrazione di probiotici, microrganismi vivi che, se ingeriti in quantità sufficienti, presentano benefici per la salute superiori a quelli dell'alimentazione di base (Mármol *et al.*, 2017; Geier *et al.*, 2006; Guidelines for the Evaluation of Probiotics in Food. FAO/WHO). I microrganismi più studiati sono appartenenti ai generi *Lactobacillus*, *Streptococcus*, *Enterococcus*, *Lactococcus* e *Bifidobacterium* (Chen *et al.*, 2012; Sah *et al.*, 2014). Questi batteri hanno effetti intraluminali, sistemici e sulla mucosa intestinale che hanno un ruolo preventivo nella cancerogenesi del colon (in particolare, competono con la flora batterica patogena, hanno attività antiossidante e pro-apoptotica, inducono la degradazione dei metaboliti cancerogeni e mutageni, salvaguardano la barriera intestinale) (Hsieh *et al.*, 2015; Molska *et al.*, 2019; Chong *et al.*, 2014).

## **4. Scopo della tesi**

Il cancro al colon-retto è una patologia multifattoriale influenzata da fattori genetici e fattori legati alla dieta e allo stile di vita. È il terzo tumore per incidenza e il secondo per mortalità al mondo, e le prospettive future prevedono un aumento significativo dei casi, soprattutto nelle nazioni in via di sviluppo. Pertanto, l'identificazione e la valutazione dei fattori di rischio e di protezione per lo sviluppo del cancro al colon-retto rappresentano obiettivi cruciali per la formulazione di strategie preventive.

È stata dimostrata una differenza, riguardante l'abbondanza ed il tipo di specie batteriche, tra il microbiota intestinale di soggetti sani e di soggetti affetti da CRC. Sono pochi i dati circa una diversa composizione del microbiota di soggetti con CRC di origine sporadica e soggetti con CRC e predisposizione genetica ai tumori.

Questo studio si propone di verificare se il microbiota intestinale sia differente in soggetti con tumori del colon di origine sporadica o mutati in soggetti con predisposizione genetica.

Il progetto rappresenta parte di un progetto più ampio, il progetto DiaMOND, "Dissecting the interaction among the intestinal Microbiota, Obesity and colon cancer Development", coordinato dalla Prof.ssa Irma Dianzani e finanziato dall'AIRC.

Lo studio ha incluso l'analisi del DNA germinale di 63 individui con polipi del colon-retto, utilizzando un pannello NGS custom di 107 geni responsabili di sindromi di predisposizione al cancro. Tale analisi permette la classificazione dei pazienti in mutati (se portatori di varianti patogeniche germinali) o sporadici. È stato inoltre effettuata l'analisi del MAM (microbiota aderente alla mucosa del polipo) di 154 pazienti tramite sequenziamento Shotgun. È stato studiato il MAM poiché i batteri a stretto contatto con gli enterociti possono svolgere un ruolo importante nella cancerogenesi del colon. A questo scopo, è stata utilizzata una nuova tecnica di prelievo del MAM tramite tampone che non compromette l'integrità dei tessuti, preservando così il polipo per le analisi istopatologiche.

## **5. Materiali e Metodi**

## **5.1 Reclutamento pazienti e raccolta campioni biologici**

I ricercatori del laboratorio di Patologia Genetica, diretto dalla Prof.ssa Irma Dianzani, hanno svolto il reclutamento dei pazienti e la raccolta dei campioni biologici presso il reparto di Gastroenterologia dell'Azienda Ospedaliera "Maggiore della Carità" di Novara. Sono stati considerati idonei i pazienti di entrambi i sessi, maggiorenni ed aventi polipi intestinali di almeno 10 mm, rimossi in sede di colonscopia. Ai pazienti è stato spiegato il progetto di ricerca a cui avrebbero partecipato e, dopo la loro autorizzazione, è stato chiesto loro di firmare il consenso informato, secondo cui acconsentono all'uso dei propri dati personali e dei campioni biologici (tampone sul polipo, sangue venoso e feci) a scopo di ricerca. Successivamente, i pazienti sono stati sottoposti ad una visita, durante la quale sono state raccolte le misure antropometriche e le informazioni relative alle abitudini alimentari, allo stile di vita e alla storia familiare per tumori.

## **5.2 Estrazione del DNA**

I campioni di sangue periferico dei pazienti vengono raccolti in provette vacutainer contenenti acido etilendiamminotetracetico (EDTA), dalle quali sono preparate almeno due aliquote da 200 µl di sangue e conservate a -20°C, fino al momento dell'utilizzo. Il DNA è estratto da sangue periferico mediante il kit QIAamp®DNA Blood Mini Kit (QIAGEN, Valencia, CA, USA), che prevede l'estrazione di DNA partendo da un volume minimo di 200 µl.

Nelle eppendorf da 1,5 ml contenente il campione di sangue di partenza vengono aggiunti 20 µl di proteasi (QIAGEN Protease) e 200 µl di Buffer AL, necessari per la degradazione proteica e per la lisi cellulare; successivamente, il campione è incubato a 56°C per 10 minuti. Al termine dell'incubazione, vengono aggiunti 200 µl di etanolo e la soluzione è trasferita in una colonna QIAamp Mini per effettuare una centrifugazione a 8000 rpm per 1 minuto, in cui il DNA si lega al filtro della colonna. Vengono poi effettuati due lavaggi: il primo con 500 µl di Buffer AW1, seguito da una centrifugazione a 8000 rpm per 1 minuto, e il secondo con 500 µl di Buffer AW2, seguito da una seconda centrifugazione a 14000 rpm per 3 minuti. Alla colonnina, posta in una nuova eppendorf, vengono aggiunti 200 µl di buffer AE e dopo 5 minuti a temperatura ambiente si procede con una centrifugazione di 1 minuto a 8000 rpm. Il DNA così estratto è sottoposto a quantificazione e poi conservato a -20°C.

### 5.3 Quantificazione del DNA

La quantificazione del DNA estratto è eseguita con due strumenti: NanoDrop TM 2000C e Qubit Fluorometer. Il NanoDrop TM 2000C, oltre ad una quantificazione rapida e accurata, permette di ottenere anche indicatori di purezza del campione, mentre il Qubit Fluorometer ha una sensibilità superiore rispetto al primo.

Il NanoDrop TM 2000C deve essere tarato prima dell'utilizzo usando la stessa soluzione con cui è eluito il DNA, 1 µl di Buffer AE. Successivamente, viene posto 1 µl di campione sul piedistallo dello strumento che viene chiuso e, tramite tensione superficiale, il campione viene mantenuto tra le due fibre ottiche. Vengono quindi eseguite 2 misurazioni:

- misurazione dell'assorbanza a 260 nm, che permette di ricavare la concentrazione del campione;
- misurazione della purezza del DNA tramite il rapporto di assorbanze 260/280, che deve essere circa 1,8 (se inferiore a 1,7 indica contaminazione da proteine, se superiore a 1,9 indica contaminazione da RNA), e il rapporto 260/230, che deve essere circa 2 (se inferiore a 1,8 indica contaminazione da etanolo o altri contaminanti).

Il Qubit dsDNA BR Assay Kit permette di analizzare campioni con concentrazioni tra 0.2 e 2000 ng/µl ed è selettivo per il DNA a doppio filamento (dsDNA). Per eseguire la quantificazione è necessario preparare una master mix composta da: un reagente fluorescente (Component A) che lega il DNA e un Buffer (Component B). Il rapporto Component A: Component B deve essere 1:200. Inoltre, devono essere preparati due standard a concentrazione nota, rispettivamente di 0 ng/µl e 100 ng/µl, utilizzando 190 µl della master mix e 10 µl di standard. La preparazione dei campioni da analizzare prevede la diluizione di 2 µl di DNA in 198 µl di master mix per ogni campione. Per effettuare la lettura al Qubit Fluorometer è necessario leggere prima gli standard per tarare lo strumento e successivamente è possibile svolgere la lettura dei campioni.

#### 5.4 Sequenziamento del DNA germinale mediante Next Generation Sequencing (NGS)

Per questo studio, il DNA germinale dei pazienti è sequenziato tramite NGS su piattaforma MiSeq (Illumina) utilizzando un pannello custom, contenente 107 geni di interesse associati a sindromi ad alta incidenza di cancro. Le library di DNA per l'NGS sono preparate utilizzando i protocolli SureSelect XT Low Input Enzymatic Fragmentation Kit e SureSelectXT HS Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library Protocol (Agilent Technologies Inc, Santa Clara, CA, USA.). Il campione deve avere una concentrazione iniziale compresa tra 10 e 200 ng/μl. Innanzitutto, per consentire l'amplificazione delle regioni di interesse, viene effettuata la frammentazione enzimatica del DNA, usando la seguente master mix (Tabella 1):

*Tabella 1: SureSelect XT Low Input Enzymatic Fragmentation Kit*

5XSureSelect Fragmentation Buffer	2 μl
SureSelect Fragmentation Enzyme	1 μl
DNA iniziale	7 μl
<b>TOTALE</b>	<b>10 μl</b>

Il programma termico utilizzato per la frammentazione è (Tabella 2):

*Tabella 2: Programma termico per la frammentazione*

Step 1	1 ciclo	37°C	15'
Step 2	1 ciclo	65°C	5'
Step 3	1 ciclo	4°C	Hold

Terminato il programma termico i campioni sono spostati in ghiaccio, ed a essi vengono aggiunti 40 μl di acqua nucleasi-free, portando così il volume a 50 μl per campione. Si procede alla preparazione della library per il successivo sequenziamento.

La procedura inizia con la preparazione della End Repair/dA-Tailing master mix (Tabella 3), per il riparo delle estremità e per l'aggiunta di code dA, utili per la successiva ligazione, in quanto l'adattatore utilizzato ha una T come prima base. La End Repair/dA-Tailing master mix è composta da:

Tabella 3: La End Repair/dA-Tailing master mix

End Repair-A Tailing Buffer	16 $\mu$ l
End Repair-A Tailing Enzyme Mix	4 $\mu$ l
<b>TOTALE</b>	<b>20 <math>\mu</math>l</b>

Per ogni campione (50  $\mu$ l) vengono aggiunti 20  $\mu$ l di master mix, e si procede alla riparazione delle estremità utilizzando il seguente programma termico:

Tabella 4: Programma termico per la riparazione delle estremità

Step 1	1 ciclo	20°C	15'
Step 2	1 ciclo	72°C	15'
Step 3	1 ciclo	4°C	Hold

A questo punto viene aggiunta la Ligation master mix, necessaria per la ligazione dell'adattatore al campione, composta come da (Tabella 5):

Tabella 5: Ligation master mix

Ligation Buffer	23 $\mu$ l
T4 DNA Ligase	2 $\mu$ l
<b>TOTALE</b>	<b>25 <math>\mu</math>l</b>

Al campione sono quindi aggiunti i 25  $\mu$ l della Ligation master mix e 5  $\mu$ l di Adaptor Oligo Mix.

Si procede con la ligazione tra campioni e adattatori secondo il seguente programma termico (Tabella 6):

Tabella 6: Programma termico per la ligazione

Step 1	1 ciclo	20°C	30'
Step 2	1 ciclo	4°C	Hold

Al termine si procede con la purificazione del DNA a cui è stato legato l'adattatore mediante l'uso delle biglie magnetiche AMPure XP, in grado di legare il DNA in maniera reversibile. Le biglie devono stare a temperatura ambiente per almeno 30 minuti prima del loro utilizzo.

Vengono aggiunti 80 µl di biglie al campione, pipettando numerose volte in modo da rendere la soluzione omogenea. Il campione viene incubato a temperatura ambiente per 5 minuti su un supporto magnetico in grado di attrarre le biglie, permettendo così di recuperare e scartare la fase liquida, trattenendo le biglie nella strip.

Vengono eseguiti due lavaggi con etanolo al 70% preparato in giornata, al termine dei quali i campioni vengono messi nel termociclatore, precedentemente portato alla temperatura costante di 37 °C, al fine di eliminare i residui di etanolo. Quanto l'etanolo è evaporato, vengono aggiunti 35 µl di acqua nucleasi-free per l'eluizione.

La strip con i campioni viene messa nel separatore magnetico, per permettere il recupero della fase liquida contenente il DNA e per poter scartare le biglie.

Prima di procedere con l'amplificazione del DNA è fondamentale assegnare ad ogni campione un indice specifico, in modo da identificare ciascun campione in modo univoco. Gli indici, forniti dal kit, oltre a contenere sequenze diverse di 8 paia di basi che fungono da tag per ogni campione, ricoprono la funzione di primer reverse.

Per l'amplificazione pre-capture, viene preparata una mix formata da (Tabella 7):

*Tabella 7: Mix per amplificazione pre-capture*

5x Herculase II Reaction Buffer	10 µl
100 mM dNTP Mix	0,5 µl
Primer forward	2 µl
Herculase II Fusion DNA Polymerase	1 µl
<b>TOTALE</b>	<b>13,5 µl</b>

Per ogni reazione sono aggiunti 13,5 µl di master mix e 2 µl di SureSelect XT HS Index Primer (uno diverso per ogni campione).

Si prosegue con l'amplificazione utilizzando il programma termico per la PCR pre-Capture, nel quale il numero dei cicli del secondo step può variare da 8 a 14 in base alla quantità del DNA (in questo caso è stato impostato a 9 cicli poiché si parte da una concentrazione iniziale di circa 50ng/µl di DNA per ogni campione) (Tabella 8):

Tabella 8: programma termico per la PCR pre-Capture

Step 1	1 ciclo	98°C	2'
Step 2	9 cicli	98°C	30''
		60°C	30''
		72°C	1'
Step 3	1 ciclo	72°C	5'
Step 4	1 ciclo	4°C	Hold

Il DNA viene purificato utilizzando le biglie magnetiche AMPure XP secondo il protocollo sopra descritto.

Il purificato viene quindi analizzato tramite lo strumento Bioanalyzer 2100 (Agilent Technologies Inc.), che permette di verificare l'avvenuta frammentazione del DNA e di misurare la dimensione e la concentrazione dei frammenti tramite elettroforesi capillare. Viene preparato un chip, che viene inserito all'interno dello strumento, utilizzando il protocollo Agilent DNA 1000 Kit (Agilent Technologies Inc.).

È preparato il Gel-Dye Mix, composto a partire da 25 µl di DNA dye concentrate aggiunto al DNA gel matrix e centrifugando per 15 minuti a 2300g la mix. 9 µl di Gel-dye mix sono inseriti nel chip, nel pozzetto contrassegnato da una G in un cerchio nero, e pressurizzati tramite l'apposita siringa. Altri 9 µl di Gel-Dye Mix vengono inseriti nei 2 pozzetti segnati da una G, mentre 5 µl di marker vengono aggiunti in tutti gli altri pozzetti del chip. 1 µl di DNA ladder è aggiunto nel pozzetto segnato da una scala e 1 µl di DNA campione negli altri pozzetti 12 pozzetti. Il chip pronto viene posto sul vortex a 2400 rpm per 1 minuto, per poi essere inserito nel Bioanalyzer. Ciò che si osserva al termine della misurazione è un elettroferogramma di ogni campione, con il peso molecolare del frammento sull'asse delle x e la relativa intensità sull'asse delle y. Inoltre, si può ricavare la concentrazione del campione attraverso l'analisi dell'area sottesa alla curva dell'elettroferogramma. Viene anche mostrata un'elettroforesi virtuale per ogni campione.



Figura 3. Chip DNA1000

Segue l'ibridazione del DNA frammentato alle capture libraries, ovvero sonde biotinilate usate nella fase di cattura per legare solo i frammenti di interesse. La fase di ibridazione necessita dai 500 ai 1000 ng di DNA in un volume di 12  $\mu$ l. Il DNA del campione viene inserito nelle strip e il volume viene portato a 12  $\mu$ l utilizzando acqua nucleasi-free; vengono aggiunti 5  $\mu$ l di SureSelect XT HS e 5  $\mu$ l di XT Low Input Blocker Mix e si esegue il seguente programma termico (Tabella 9):

*Tabella 9: Programma Termico per l'ibridazione*

Step 1	1 ciclo	95°C	5'
Step 2	1 ciclo	65°C	10'
Step 3	1 ciclo	65°C	1'
Step 4	60 cicli	65°C	1'
		37°C	3''
Step 5	1 ciclo	65°C	Hold

Durante il terzo step è necessario mettere in pausa il termociclatore, per aggiungere 13  $\mu$ l di Capture Library Hybridization Mix composta da (Tabella 10):

*Tabella 10: Capture Library Hybridization Mix*

Soluzione al 25% di SureSelect RNase Block	2 $\mu$ l
Capture Library <3 Mb	2 $\mu$ l
SureSelect Fast Hybridization Buffer	6 $\mu$ l
Acqua nucleasi-free	3 $\mu$ l
<b>TOTALE</b>	<b>13 <math>\mu</math>l</b>

Al termine dell'ibridazione viene effettuata la cattura del DNA con le sonde ibridate utilizzando le biglie ricoperte di streptavidina (Dynabeads MyOne Streptavidin T1 magnetic beads, ThermoFisher). Le biglie di streptavidina si legano, per l'interazione biotina-streptavidina, alle sonde biotinilate, che sono complementari al DNA target.

Le biglie devono essere preparate prima della fine dell'ibridazione attraverso tre lavaggi con 200  $\mu$ l di SureSelect Binding Buffer, per eliminare il liquido di conservazione e scoprire i siti di legame per la biotina.

La fase di cattura prevede il trasferimento di tutto il prodotto di ibridazione nelle strip contenenti le biglie lavate e la seguente incubazione in un mixer a 1500 rpm per 30 minuti.

Successivamente, le strip vengono messe in un separatore magnetico per poter recuperare ed eliminare la fase liquida, lasciando nelle strip solo le biglie, che vengono risospese in 200 µl di Wash Buffer 1.

Le strip vengono messe nuovamente nel separatore magnetico e il Wash Buffer 1 viene eliminato. A questo seguono 6 lavaggi con Wash Buffer 2 preriscaldato a 70°C, e al termine dell'ultimo lavaggio le biglie vengono risospese in 25 µl di acqua nucleasi-free, per procedere alla fase di amplificazione post-capture. Quest'ultima necessita della preparazione di una PCR master mix così composta (Tabella 11):

*Tabella 11: PCR master mix*

Acqua nucleasi-free	12,5 µl
5x Herculase II Reaction Buffer	10 µl
Herculase II Fusion DNA Polymerase	1 µl
100mM dNTP Mix	0,5 µl
SureSelect Post-Capture Primer Mix	1 µl
<b>TOTALE</b>	<b>25 µl</b>

All'interno di ogni campione sono aggiunti i 25 µl di tale mix, per poi procedere con il seguente programma termico nel termociclatore:

*Tabella 12: Programma termico per PCR*

Step 1	1 ciclo	98°C	2'
Step 2	12 cicli	98°C	30''
		60°C	30''
		72°C	1'
Step 3	1 ciclo	72°C	5'
Step 4	1 ciclo	4°C	Hold

Al termine dell'amplificazione le strip sono posizionate nel separatore magnetico, rimuovendo le biglie di streptavidina e recuperando la fase liquida, che viene pipettata in nuove strip. A questo punto si effettua la purificazione della libreria dai reagenti utilizzati per l'amplificazione, utilizzando le biglie magnetiche AMPure XP.

Infine, deve essere usato il Bioanalyzer 2100, per misurare la concentrazione e valutare la frammentazione dei campioni ottenuti, utilizzando il kit High Sensitivity, adatto per concentrazioni minori rispetto al protocollo DNA1000.

Ottenute le librerie di DNA di ogni paziente, per procedere al sequenziamento, si crea un pool di tutti i campioni, che deve essere di 20  $\mu$ l con concentrazione 10 nM. Si preleva quindi un volume appropriato di ogni libreria di DNA, in quanto all'interno del pool tutti i campioni devono essere presenti in concentrazioni equimolari. Se il volume finale ottenuto è superiore a 20  $\mu$ l, è necessario utilizzare lo speed vacuum concentrator per circa 45 minuti, in modo che il liquido evapori e si possa risospendere il pool in 20  $\mu$ l di acqua nucleasi-free.

Si misura al Qubit la concentrazione del pool a 10 nM per poi diluirlo a 4 nM e misurando nuovamente la concentrazione. In entrambi i casi per la misurazione si utilizza il kit High Sensitivity. A questo punto si denaturano 5  $\mu$ l di pool utilizzando 5  $\mu$ l di NaOH a concentrazione 0,2 N (preparato in precedenza dal NaOH 10 N), e in seguito si aggiungono 990  $\mu$ l di Buffer HT1 (Hybridization Buffer fornito dal kit). Si porta il pool dalla concentrazione di 4 nM alla concentrazione di 9 pM (270  $\mu$ l del pool denaturato + 330  $\mu$ l del Buffer HT1), concentrazione consigliata per il sequenziamento su MiSeq con una cartuccia v2 (Illumina).

Parallelamente è preparato il Phix control, ovvero il DNA di controllo della corsa, diluendo 2  $\mu$ l di Phix control (Illumina) in 3  $\mu$ l di Tris-Cl pH 8,5 con 0,1% Tween 20, ottenendo il Phix a concentrazione 4 nM. Questo è denaturato con 5  $\mu$ l di NaOH 0,2 N e diluito in 990  $\mu$ l di Buffer HT1 per ottenere una concentrazione di 20 pM. A questa soluzione sono aggiunti dai 375  $\mu$ l ai 225  $\mu$ l di HT1 per ottenere una diluizione di Phix concentrata 12,5 pM. Infine, vengono aggiunti 6  $\mu$ l di Phix 12,5 pM al pool a 9 pM, che è in seguito caricato nella cartuccia, nel pozzetto apposito.

La cartuccia è inserita nel MiSeq, insieme al Buffer e alla flow cell, su cui verranno generati i cluster e verrà effettuato il sequenziamento. La reazione di sequenziamento avviene a seguito di un'amplificazione "a ponte" dei campioni legati alla flow cell per formare i cluster. Dopo essere stati generati, i cluster vengono sottoposti a imaging, utilizzando LED e combinazioni di filtri specifici per ciascuno dei quattro dideossinucleotidi con marcatura fluorescente. Questo processo si ripete ad ogni ciclo di sequenziamento. Dopo l'analisi delle immagini, il software esegue l'identificazione delle basi, il filtraggio e il calcolo dei punteggi qualitativi.

## 5.5 Analisi *in silico*

Le varianti nonsense, missense e frameshift emerse da NGS sono filtrate per frequenza selezionando MAF (Minor Allele Frequency) <0.01 nei database 1000 Genomi EUR, ExAc NFE EUR, poiché le varianti patogenetiche sono rare. Viene poi ricercata la patogenicità delle varianti trovate sui database:

- ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>),
- Varsome (<https://varsome.com/>),
- Franklin (<https://franklin.genoox.com/clinical-db/home>).

Per ogni variante identificata come patogena, viene consultato il sito Ensembl genome browser (<https://www.ensembl.org>). Si procede infine con la conferma tramite sequenziamento in Sanger.

## 5.6 Disegno Primer

Per le regioni geniche in cui sono state identificate mutazioni con NGS e per le quali si procede alla validazione, sono stati disegnati primers utilizzando come riferimento le sequenze disponibili nel database Ensembl (<https://www.ensembl.org/index.html>).

Le caratteristiche ottimali dei primers sono:

- Lunghezza di 18-24 nucleotidi;
- Percentuale di GC del 40-60%;
- T di annealing di 55°-65°C.

Nella tabella 13 sono riportati i primers utilizzati per l'amplificazione tramite PCR e il sequenziamento con il metodo di Sanger dei geni studiati:

Tabella 13: Primers utilizzati

<b>NTHL1_exon2F</b>	5'- CACCCCTACCCTACCTTCAC - 3'
<b>NTHL1_exon2R</b>	5'- TTTTGTCCTCATGGCACGG- 3'
<b>MUTYH_exon13F</b>	5'- AGAGAATGGAGGGAATCGGC- 3'
<b>MUTYH_exon13R</b>	5'- TGCTCACTTACCTCCCAAG- 3'
<b>MSH6_exon4F</b>	5'- GAGTCAGAACCACCCAGACA - 3'
<b>MSH6_exon4Rex</b>	5'- AAGGCTGTATCCCATCGGTT 3'
<b>FANCE_exon5F</b>	5'- CTTGAGACAGCCTAGCAGAG - 3'
<b>FANCE_exon5R</b>	5'- GCACTCTGAGTACCTCCCAA - 3'

<b>PRF1_exon6F</b>	5'- GAGGTCTCTCTCTTCTCGCAG - 3'
<b>PRF1_exon6R</b>	5'- TAGGTTTGGTGGGAAGGAGGC - 3'

## 5.7 PCR

Le reazioni di PCR sono allestite utilizzando una coppia di oligonucleotidi (forward, reverse), e per ogni coppia di primers sono state effettuate diverse prove al fine di trovare le condizioni ottimali di amplificazione, relativamente alla temperatura di annealing ed alla concentrazione di cloruro di magnesio. Nella tabella 14 sottostante sono riportate le condizioni ottimali di temperatura di annealing e concentrazione di MgCl<sub>2</sub> per favorire l'azione della polimerasi.

*Tabella 14: Condizioni ottimali di temperatura di annealing e concentrazione di MgCl<sub>2</sub>*

<b>Gene</b>	<b>Concentrazione MgCl<sub>2</sub></b>	<b>Temperatura annealing</b>
<b>NTHL1_exon2</b>	1mM	60°C
<b>MUTYH_exon13</b>	1mM	59°C
<b>MASH6_exon4</b>	1 mM	59°C
<b>FANCE_exon5</b>	1mM	57°C
<b>PRF1_exon6</b>	1mM	60°C

Il DNA dei pazienti è stato amplificato tramite PCR (Polymerase Chain Reaction) in un volume finale totale di 15 µl. La master mix è composta da (Tabella 15):

*Tabella 15: Master Mix PCR*

<b>Primer forward 10µM</b>	0,75 µl
<b>Primer reverse 10µM</b>	0,75 µl
<b>MgCl<sub>2</sub> 25mM</b>	0,6 µl
<b>GoTaq Flexi Buffer 5x</b>	3 µl
<b>dNTPs 10 mM</b>	0,3 µl
<b>GoTaq® DNA FlexiPolymerase 5U/l</b>	0,15 µl
<b>TOTALE</b>	<b>13 µl</b>

Ad ogni mix è stata aggiunta H<sub>2</sub>O per raggiungere un volume di 13 µl. In seguito, sono stati aggiunti 2 µl di DNA di controllo nel controllo positivo, 2 µl di acqua nel controllo negativo e 2 µl di DNA a concentrazione iniziale di 15 ng/µl da amplificare nel pozzetto relativo al paziente. Le reazioni sono poi inserite nel termociclatore con il programma termico:

*Tabella 16: Programma termico per PCR*

Step 1	1 ciclo	95°C	2'
Step 2	35 cicli	95°C	30''
		Ta (Tabella 14)	30''
		72°C	45''
Step 3	1 ciclo	72°C	2'
Step 4	1 ciclo	4°C	Hold

## 5.8 Elettroforesi

Viene effettuata l'elettroforesi su gel di agarosio al 2% dei prodotti di PCR per verificare l'avvenuta amplificazione.

Per preparare il gel è necessario sciogliere 2 g di agarosio in 100 ml di TAE 1x (Tris-Acetate-EDTA) in una beuta, miscelare bene e scaldare con il microonde fino a che la soluzione non arriva ad ebollizione. È necessario che la soluzione si raffreddi per aggiungere 4 µl di Gel Red<sup>TM</sup> Nucleic Acid Stain, un colorante fluorescente in grado di intercalarsi tra le basi del DNA. La soluzione è poi versata in una vaschetta munita di pettinini che permettono la formazione dei pozzetti nel gel. Il gel solidificato viene posto nella camera elettroforetica e immerso nel TAE 1x.

Per eseguire l'elettroforesi è necessario preparare una mix composta da 8 µl di H<sub>2</sub>O e 2 µl di Orange 6x (Orange 1 mg/ml, glicerolo 50%: il colorante è utilizzato per facilitare il caricamento del DNA nel pozzetto ed è utile per tracciare la migrazione del campione, mentre il glicerolo permette di aumentare la densità del campione per impedirne la dispersione nel tampone). A questa mix vengono aggiunti 2 µl di campione e viene caricata all'interno dei pozzetti del gel. In un pozzetto vengono messi 2 µl di marker di corsa FastGene® 100 bp DNA Ladder RTU, NIPPON Genetics Europe GmbH. Al termine dell'elettroforesi il gel è portato al transilluminatore a raggi UV Biorad Model ChemiDoc<sup>TM</sup> Touch Imaging System, che permette di osservare le bande che rappresentano i frammenti amplificati. Nel controllo

negativo di ogni PCR non devono essere presenti bande che, altrimenti, indicherebbero una contaminazione.

### 5.9 Purificazione dell'amplificato

Il prodotto di amplificazione viene purificato per eliminare i reagenti della PCR. Viene utilizzato il kit di purificazione EuroSAP PCR Enzymatic Clean-up kit (EuroClone® S.p.A), contenente due enzimi: Exonuclease 1 (Exo1) e Shrimp Alkaline Phosphatase (SAP). Exo1 degrada i primers a singola elica di DNA e SAP defosforila i nucleotidi.

Viene preparata una master mix composta da (Tabella 17):

Tabella 17: Master Mix Purificazione dell'amplificato

SAP (2U/ $\mu$ l)	0,5 $\mu$ l
Exo1 (10U/ $\mu$ l)	0,5 $\mu$ l
H <sub>2</sub> O	3 $\mu$ l
<b>TOTALE</b>	<b>4 <math>\mu</math>l</b>

In ogni pozzetto sono messi 4  $\mu$ l di mix e 2  $\mu$ l di prodotto di PCR. Infine, i campioni sono inseriti nel termociclatore con il seguente programma termico (Tabella 18):

Tabella 18: Programma Termico per purificazione amplificato

Step 1	1 ciclo	37°C	15'
Step 2	1 ciclo	80°C	15'
Step 3	1 ciclo	4°C	Hold

## 5.10 Preparazione delle reazioni di sequenza

Per eseguire le reazioni di sequenza è utilizzato il BigDYE® Terminator V1.1 CycleSequencing Kit. Viene composta una mix come segue (Tabella 19):

Tabella 19 : BigDYE® Terminator V1.1 CycleSequencing Kit

<b>Purificato PCR</b>	2 µl
<b>Primer forward 3,2 µM</b>	1 µl
<b>Primer reverse 3,2 µM</b>	1 µl
<b>Big Dye</b>	1,2 µl
<b>H<sub>2</sub>O</b>	5,8 µl

Questo metodo di sequenziamento sfrutta la capacità della polimerasi di incorporare al 3' della catena di DNA in sintesi sia dideossinucleotidi, coniugati con fluorocromi, sia deossinucleotidi non marcati. L'incorporazione di dideossinucleotidi provoca il blocco della sintesi di DNA.

Avremo quindi dei frammenti di varia lunghezza che terminano con una base specifica. Essendo i dideossinucleotidi marcati con diversi fluorocromi, ad ogni diversa base di terminazione corrisponde una fluorescenza diversa, che verrà letta tramite elettroforesi capillare.

Le strip vengono inserite nel termociclatore con il programma termico descritto di seguito (Tabella 20):

Tabella 20: Programma termico per reazioni di sequenza

Step 1	1 ciclo	96°C	1'
Step 2	25 cicli	95°C	15''
		50°C	5''
Step 3	1 ciclo	60°C	4''
Step 4	1 ciclo	4°C	Hold

### **5.11 Purificazione delle reazioni di sequenza**

Le reazioni di sequenza vengono poi purificate attraverso la cromatografia a gel filtrazione, utilizzando il kit Performa Spin Columns (Edge Bio).

Il primo passaggio prevede la preparazione delle colonnine con il Sephadex G-50 (Sigma-Aldrich), sostanza costituita da sfere molto piccole ricche di gruppi idrossilici, utilizzato come fase stazionaria. Aggiungendo ad esso 700 µl di acqua e aspettando 15 minuti con tappo aperto e 15 minuti con tappo chiuso, le sfere, gonfiandosi, producono il gel utilizzato per la filtrazione. La matrice così prodotta è in grado di recuperare frammenti di DNA con dimensione superiore a 16 bp, e di rimuovere il 98% di sali, nucleotidi e impurità a basso peso molecolare.

Al termine dell'attesa le colonnine vengono centrifugate per 3 minuti a 850 rcf per rimuovere l'acqua in eccesso e viene aggiunto sulla colonnina tutto il prodotto della reazione di sequenza, facendo attenzione a non toccare il gel. A ciò segue una seconda centrifugazione a 850 rcf per 3 minuti, che permette di recuperare la reazione di sequenza purificata.

### **5.12 Sequenziamento di Sanger**

In seguito alla purificazione delle reazioni di sequenza viene eseguita l'analisi delle sequenze utilizzando il sequenziatore SeqStudio Genetic Analyzer System (Thermo Fisher).

Viene utilizzata una piastra ottica da 96 pozzetti, e in ciascun pozzetto sono caricati 10 µl di formammide, necessaria per la denaturazione del DNA, e 5 µl di reazioni di sequenza purificate. Le sequenze sono denaturate per 2 minuti a 96°C, per poi procedere con il sequenziamento.

### **5.13 Analisi delle sequenze**

Al termine del sequenziamento, i risultati vengono analizzati allineando le sequenze ottenute con le sequenze wild-type di riferimento utilizzando EMBOSS Needle ([https://www.ebi.ac.uk/Tools/psa/emboss\\_needle/](https://www.ebi.ac.uk/Tools/psa/emboss_needle/)), e valutando tramite il programma Chromas gli elettroferogrammi delle sequenze analizzate (<https://technelysium.com.au/wp/chromas/>).

### **5.14 Conferma delle varianti sul DNA tumorale**

Per ogni variante trovata tramite NGS e poi confermata in Sanger sul DNA germinale, viene eseguita la ricerca della mutazione su DNA estratto da tessuto tumorale dal reparto di Anatomia Patologica dell'Azienda Ospedaliera "Maggiore della Carità" di Novara (se si ha disponibilità del materiale) come descritto nel capitolo 5.2 con le stesse metodiche utilizzate per il DNA germinale sopra riportate.

### **5.15 Classificazione pazienti**

A seguito della conferma delle varianti in Sanger, i pazienti sono stati classificati in mutati (se portatori di varianti patogeniche germinali), familiari (se rientranti nei criteri per la familiarità che indicano una sospetta predisposizione genetica, ma senza varianti patogeniche germinali identificate) o sporadici (se non portatori di varianti germinali e non rientranti nei criteri per la familiarità). La familiarità viene definita in base ai seguenti criteri:

- età di esordio <50 anni,
- diagnosi di un altro tumore maligno istologicamente diverso in aggiunta ai polipi coloretali,
- presenza di due parenti di primo o secondo grado affetti da cancro nella stessa discendenza,
- presenza di almeno un parente di primo grado affetto da CRC (Syngal *et al.*, 2015; Gupta *et al.*, 2020).

### **5.16 Analisi del microbioma adeso al polipo e luminale (Shotgun Sequencing)**

Le analisi del microbiota sono svolte grazie alla collaborazione con il Centro per le malattie autoimmuni e allergiche (CAAD) di Novara.

Il microbiota adeso al polipo è prelevato tramite un tampone eNat FLOQSwabs® (COPAN, Brescia, Italy) e posto in liquido di conservazione. Il tampone è eseguito prima che il polipo venga fissato per la successiva analisi anatomopatologica, condotta presso il reparto di Anatomia Patologica dell'Azienda Ospedaliera "Maggiore della Carità" di Novara diretto dal Professor Renzo Boldorini.

L'estrazione del DNA microbico adeso al polipo è effettuata con il kit QIAamp DNA Microbiome (Qiagen), che utilizza una combinazione di lisi chimica e meccanica del campione. Per la lisi chimica il campione viene omogenizzato con un buffer specifico, mentre per la lisi meccanica sono utilizzate PowerBead Pro Tubes contenenti biglie di vetro.

Il sequenziamento metagenomico Shotgun è eseguito dalla ditta Novogene (UK) Company Limited. Le librerie sono state preparate con il kit NexteraXT DNA Library Preparation Kit (Illumina) e sequenziate sulla piattaforma Illumina NovaSeq.

Il profilo tassonomico, con la stima degli arricchimenti relativi a livello di specie, è stato eseguito utilizzando MetaPhlAn versione 4.0, con la versione del database dei marcatori 201901. L'analisi statistica a valle è stata eseguita dal nostro collaboratore Dr. Edoardo Pasolli attraverso script personalizzati scritti nell'ambiente R. La dimensione dell'effetto di analisi discriminante lineare (LEfSe, linear discriminant analysis, effect size) è stata utilizzata per

identificare i taxa del MAM che sono significativamente diversi tra i gruppi mutati e sporadici o familiari/mutati e sporadici ( $p < 0.05$ ).

## **6.Risultati**

## 6.1 Risultati del sequenziamento del DNA germinale mediante NGS

Durante il periodo di tirocinio sono state svolte le analisi NGS sul DNA germinale di 63 pazienti con polipi del colon-retto. In media, il 77,6% delle reads copre le regioni target. La maggior parte dei campioni presenta il 99,25% delle reads con un coverage maggiore del 10x, mentre la copertura media è 120x. Dall'analisi sono emerse 25653 varianti germinali, da cui sono state escluse quelle con MAF (minor allele frequency)  $\geq 0.01$ , dal momento che le varianti patogeniche sono rare, basandosi sui database 1000Genomi EUR e ExAc NFE EUR. Delle 1542 varianti ottenute, 163 erano varianti frameshift, nonsense e missense. Tra esse, 5 varianti sono state considerate patogeniche e 1 probabilmente patogenica (Tabella 21) secondo i database:

- Clinvar (<https://www.ncbi.nlm.nih.gov/clinvar/>)
- Varsome (<https://varsome.com/>)
- Franklin (<https://franklin.genoox.com/clinical-db/home>)

I risultati del sequenziamento NGS sono stati visualizzati tramite il software IGV (Integrative Genomics Viewer, <https://igv.org/app/>).

Tabella 21: varianti patogeniche germinali classificate come patogeniche o probabilmente patogeniche

ID paziente	Gene	cDNA	Proteina	Sindrome associata al gene
s129	<i>TP53</i>	c.329G>T	p.Arg110Leu	Sindrome di Li Fraumeni (autosomica dominante)
s219	<i>PRF1</i>	c.695G>A	p.Arg232His	Linfoistiocitosi emofagocitica familiare (autosomica recessiva)
s242	<i>MUTYH</i>	c.1103G>A	p.Gly368Asp	Poliposi associata a MUTYH (autosomica recessiva)
s243	<i>FANCE</i>	c.1111C>T	p.Arg371Trp	Anemia di Fanconi (autosomica recessiva)
s293	<i>MSH6</i>	c.2570A>C	p.Asp857Ala	Sindrome di Lynch (autosomica dominante)
s297	<i>NTHL1</i>	c.244C>T	p.Gln82*	Poliposi associata a NTHL1 (autosomica recessiva)

I pazienti sono eterozigoti per le varianti riportate nella tabella 21.

Nel paziente *s129* è stata riscontrata la variante p.Arg110Leu del gene *TP53*, considerata patogenica da tutti i siti di predizione. *TP53* è un oncosoppressore che svolge una funzione fondamentale nella protezione dell'organismo dallo sviluppo del cancro. La proteina codificata da questo gene, p53, regola la proliferazione cellulare e l'espressione di geni proapoptotici a seguito di danno al DNA (Naccarati *et al.*, 2012). I soggetti affetti dalla sindrome di Li-Fraumeni, che hanno varianti patogeniche germinali in eterozigosi nel gene *TP53*, mostrano un aumento del rischio di sviluppare vari tipi di tumore (Naccarati *et al.*, 2012). In particolare, nel CRC la frequenza delle mutazioni somatiche in *TP53* è circa del 50% (Naccarati *et al.*, 2012). Non ci sono altre informazioni cliniche per supportare o meno che il paziente *s129* abbia la sindrome di Li-Fraumeni.

Nel paziente *s219* è stata individuata la variante p.Arg232His del gene *PRF1*. Questo gene codifica per la proteina perforina, fondamentale per l'attività dei linfociti citotossici (Voskoboinik *et al.*, 2006). Studi funzionali hanno dimostrato che questa specifica variante patogenica provoca la maturazione parziale della perforina (Risma *et al.*, 2006) e non ne permette l'espressione nei linfociti citotossici (Feldmann *et al.*, 2002; Sieni *et al.*, 2012), che presentano un'attività citotossica del 30% rispetto ai wild-type (Voskoboinik *et al.*, 2006). Le mutazioni in omozigosi, che provocano la perdita della funzione in *PRF1*, causano la linfoistiocitosi emofagocitica familiare (FHL), ma è stato ipotizzato che alcune mutazioni missense, che causano parziale perdita di espressione o funzione, portino a insorgenza tardiva di FHL e neoplasie (Chaudhry *et al.*, 2016).

Nel paziente *s242* è stata riscontrata la variante in eterozigosi p.Gly368Asp nel gene *MUTYH*, che codifica per una glicosilasi che scinde le basi dell'adenina accoppiate inappropriatamente con la guanina, citosina o 8-oxo-7,8-diidroguanina (Brinkmeyer *et al.*, 2015). Studi funzionali hanno dimostrato una ridotta attività glicosilasica e di legame al DNA per questa variante a confronto con la proteina wild-type (Ali *et al.*, 2008; Ruggieri *et al.*, 2013; Komine *et al.*, 2015; Goto *et al.*, 2010). Varianti patogeniche germinali in omozigosi in questo gene sono responsabili della poliposi *MUTYH*-associata (MAP) (Nielsen *et al.*, 1993; Rizzolo *et al.*, 2018), ma un rischio aumentato è stato riscontrato anche in pazienti con mutazioni di *MUTYH* in eterozigosi per CRC (Win *et al.*, 2011).

Nel paziente *s243* è stata rilevata la variante p.Arg371Trp nel gene *FANCE*, che impedisce l'interazione tra *FANCE* e *FANCD2* (Nookala *et al.*, 2007), fondamentale nel processo di

riparazione del DNA (Esteban-Jurado *et al.*, 2016). Mutazioni in omozigosi in *FANCE* provocano l'anemia di Fanconi (FA). I pazienti con FA sono suscettibili a vari tipi di cancro, compreso CRC (Parsa *et al.*, 2022). Studi effettuati sui genitori eterozigoti dei pazienti con FA hanno mostrato un aumentato rischio di cancro per i geni *BRCA1*, *BRCA2*, *PALB2*, *BRIP1* e *RAD51C*. Molti studi hanno riscontrato eterozigosi per mutazioni patogene in soggetti con vari tipi di cancro. Si ritiene che conferiscano un rischio moderato (McReynolds *et al.*, 2022; Mehta & Ebens, 1993). In particolare, *FANCE* è stato trovato mutato in eterozigosi a livello germinale in due fratelli con CRC (Esteban-Jurado *et al.*, 2016).

Nel paziente **s293** è stata identificata la variante p.Asp857Ala del gene *MSH6*, che non è mai stata riportata prima in letteratura, quindi non si può assumere il suo effetto patogeno. Gli strumenti di previsione *in silico* utilizzati, Franklin e Varsome, prevedono un effetto patogeno moderato per questa variante, classificandolo come un fattore di rischio moderato. *MSH6* è un gene del MMR e, se mutato, causa la sindrome di Lynch (Maratt *et al.*, 2022).

Nel paziente **s297** è stata rilevata la variante p.Gln82\* in *NTHL1*, che codifica per una glicosilasi del DNA coinvolta nella via BER (base-excision repair). Mutazioni bialleliche in *NTHL1* sono responsabili della poliposi associata a NTHL1 (NAP) (Weren *et al.*, 2015). Questa variante è considerata patogenica dai siti di predizione, ma non è ancora chiaro se i portatori in eterozigosi della variante abbiano un aumentato rischio di sviluppare il cancro al colon (Weren *et al.*, 2015).

## 6.2 Conferme delle varianti con sequenziamento di Sanger

È stata eseguita l'analisi di Sanger sul DNA germinale per confermare le varianti identificate come patogeniche da NGS. In seguito, è stata effettuata l'analisi sul tessuto tumorale per valutare il ruolo dei geni mutati nella cancerogenesi e per valutare l'eventuale perdita del secondo allele causata da mutazioni somatiche. Il DNA estratto dal tessuto tumorale è stato quindi sequenziato tramite sequenziamento di Sanger per verificare l'eventuale scomparsa dell'allele wild-type nel sito della mutazione identificata a livello germinale.

Le figure che seguono rappresentano gli elettroferogrammi relativi ai siti di mutazione sul DNA germinale e sul DNA tumorale.

### 6.2.1 DNA germinale e tumorale: paziente s129

#### *TP53 c.329G>T p.Arg110Leu*

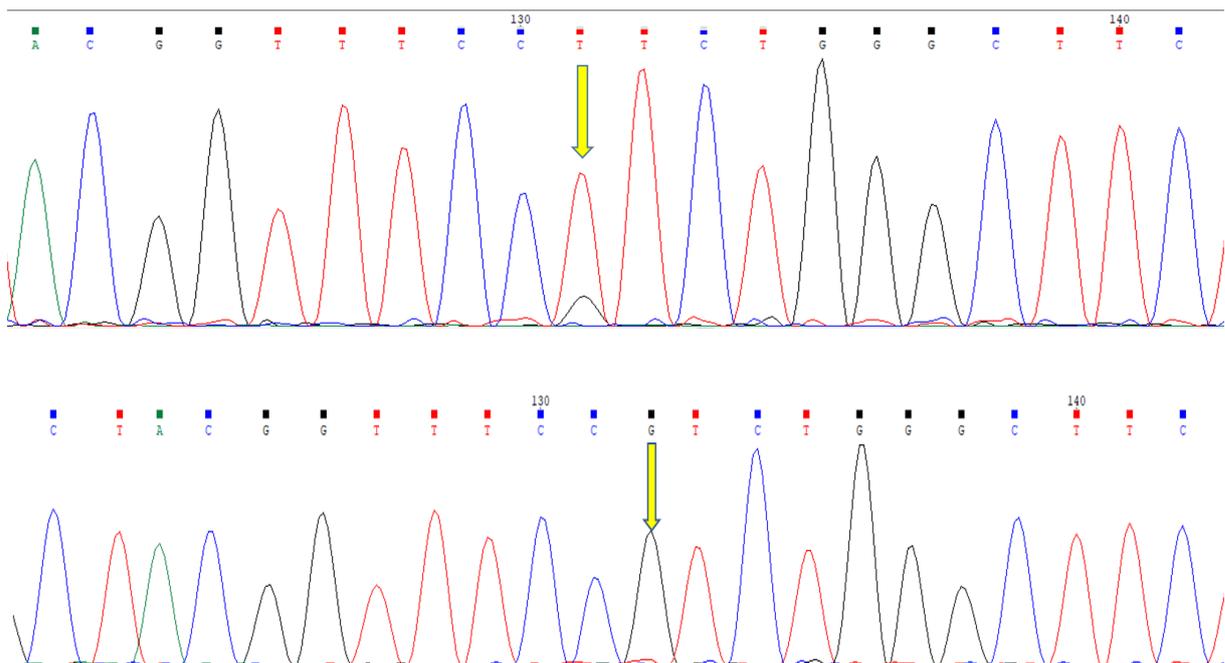


Figura 4. Elettroferogramma relativo al sequenziamento del gene TP53 del paziente s129 su DNA germinale (in alto) e DNA tumorale (in basso). La freccia evidenzia la variante c.329G>T p.Arg110Leu, che non si conferma nel tumorale.

### 6.2.2 DNA germinale e tumorale: paziente s219

#### *PRF1* c.695G>A p.Arg232His

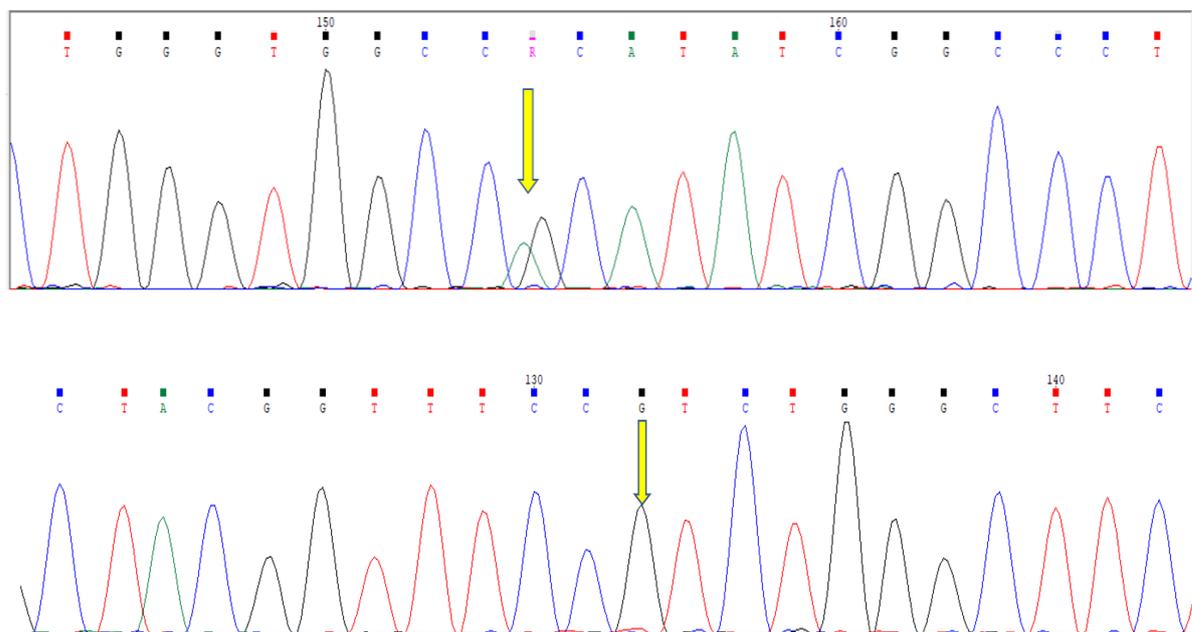


Figura 5. Elettroferogramma relativo al sequenziamento del gene *PRF1* del paziente s219 su DNA germinale (in alto) e DNA tumorale (in basso). La freccia evidenzia la variante c.695G>A p.Arg232His, che non si conferma nel tumorale.

La mancata conferma di queste varianti su DNA tumorale apre l'ipotesi di mutazioni somatiche limitate alle cellule ematiche. In letteratura sono riportati casi in cui varianti patogeniche apparentemente germinali di *TP53* nelle cellule del sangue periferico non venivano identificate sul DNA del tessuto tumorale poiché insorte come conseguenza dell'emopoiesi clonale (Mester *et al.*, 2020).

### 6.2.3 DNA germinale e tumorale: paziente s242

#### *MUTYH c.1103G>A p.Gly368Asp*

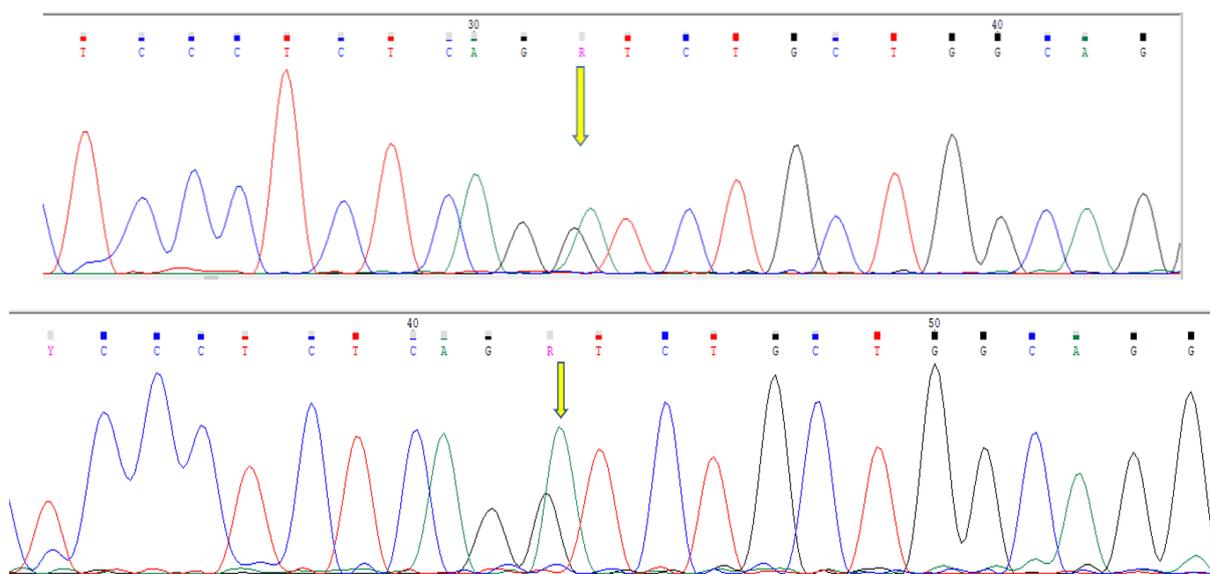


Figura 6. Elettroferogramma relativo al sequenziamento del gene *MUTYH* del paziente s242 su DNA germinale (in alto) e DNA tumorale (in basso). La freccia evidenzia la variante *c.1103G>A p.Gly368Asp*.

### 6.2.4 DNA germinale e tumorale: paziente s243

#### *FANCE c.1111C>T p.Arg371Trp*

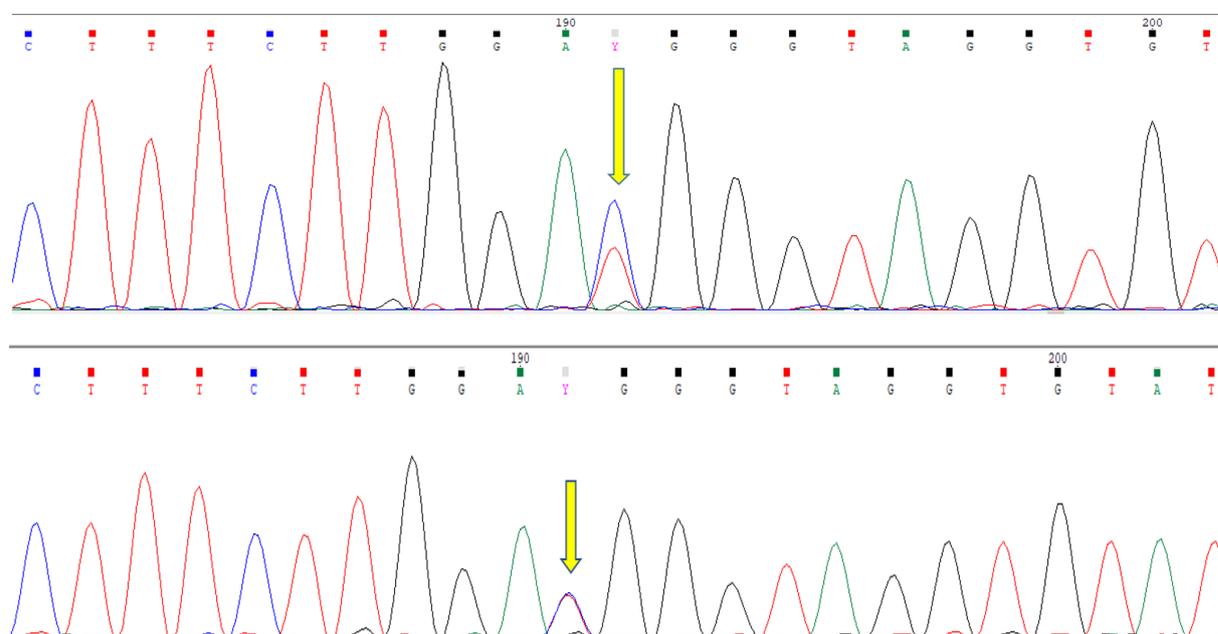


Figura 7. Elettroferogramma relativo al sequenziamento del gene *MUTYH* del paziente s243 su DNA germinale (in alto) e DNA tumorale (in basso). La freccia evidenzia la variante *c.1111C>T p.Arg371Trp*.

### 6.2.5 DNA germinale e tumorale: paziente s293

#### *MSH6 c.2570A>C p.Asp857Ala*

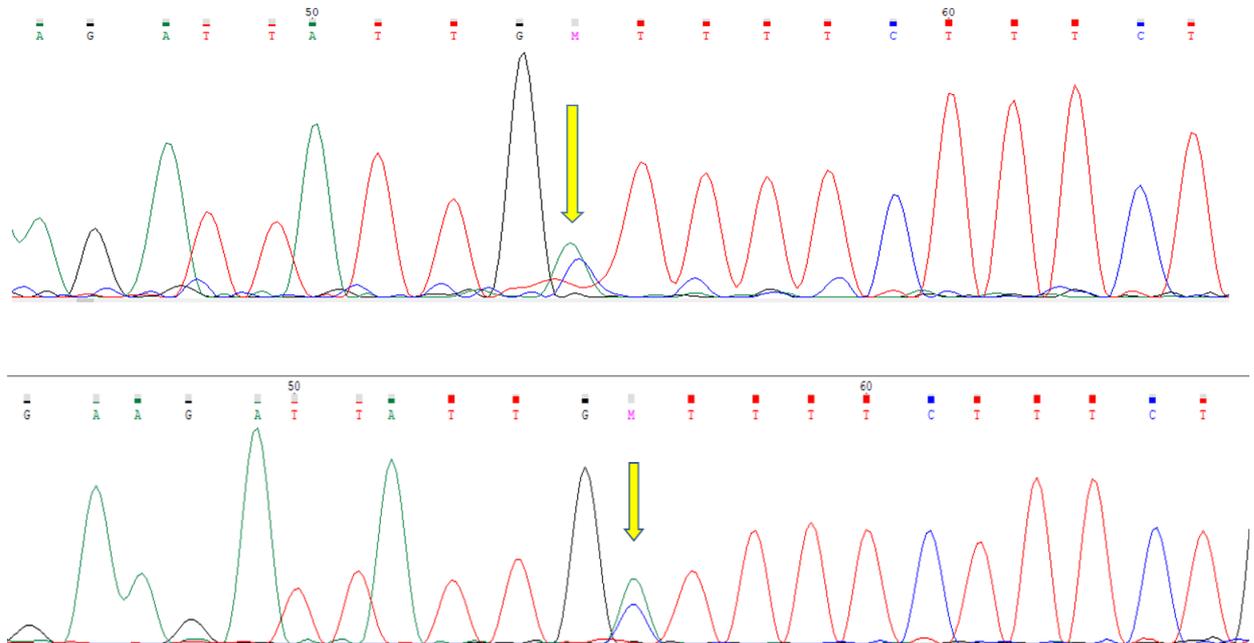


Figura 8. Elettroferogramma relativo al sequenziamento del gene *MSH6* del paziente s293 su DNA germinale (in alto) e DNA tumorale (in basso). La freccia evidenzia la variante *c.2570A>C p.Asp857Ala*.

### 6.2.6 DNA germinale e tumorale: paziente s294

#### *NTHL1 c.244C>T p.Gln82\**

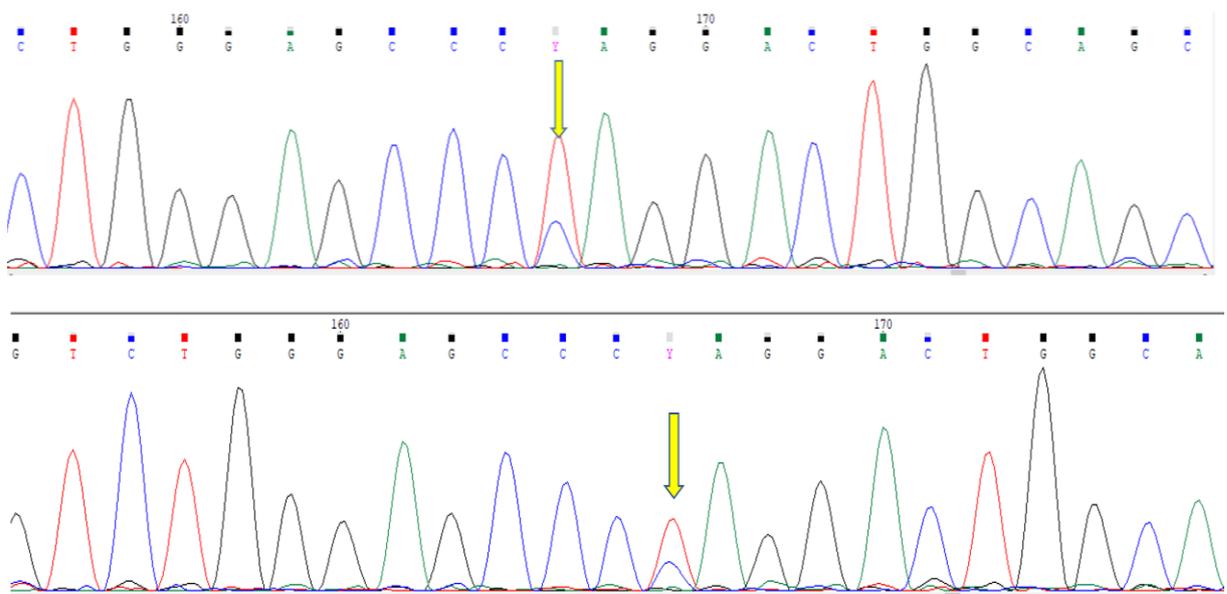


Figura 9. Elettroferogramma relativo al sequenziamento del gene *NTHL1* del paziente s297 su DNA germinale (in alto) e DNA tumorale (in basso). La freccia evidenzia la variante *c.244C>T p.Gln82\**.

Le varianti nei geni *MUTYH*, *FANCF*, *MSH6* e *NTHL1* sono state trovate in eterozigosi anche nel tumore. Questo può essere spiegato da:

- la tecnica Sanger non è abbastanza sensibile per rilevare la perdita di eterozigosità in presenza di cellule stromali wild-type. Sono in corso analisi più specifiche.
- il secondo allele potrebbe essere inattivato con meccanismi epigenetici,
- il gene potrebbe indurre predisposizione anche in eterozigosi.

### 6.3 Varianti patogeniche evidenziate nell'intera casistica dei pazienti

I pazienti analizzati durante il periodo del mio tirocinio sono stati aggiunti alla casistica già analizzata in precedenza nel laboratorio in cui ho svolto la tesi, per arrivare ad un totale di 242 pazienti.

Nella tabella 22 sottostante sono elencate le varianti identificate nell'intera casistica di pazienti analizzati nel laboratorio di Patologia Genetica. Ciascuna variante è stata classificata come variante ad alto rischio (in rosso) o a rischio moderato (in arancione) per lo sviluppo del tumore al colon-retto in base al suo effetto sulla funzionalità della proteina e alle informazioni presenti in letteratura.

Pz ID	Sesso	Età	Parenti con Cancro	Gene e variante patogenica	Funzione del gene	Sindrome ereditaria da cancro o fattore di rischio
s2	F	69	padre: cancro alla vescica; fratello: cancro epatico	<i>FANCM</i> c.4561_4562insTT p.Glu1521fs	HRR	Fattore di rischio moderato, eterozigosi per FA
s9	F	63	No	<i>GALNT12</i> c.907G>A p.Asp303Asn	O-glicosilazione	Fattore di rischio moderato
s21	M	64	padre: tumore pancreatico; fratello (s39): cancro uratrale	<i>BMPRIA</i> c.1327C>T p.Arg443Cys	Regolazione della segnalazione di WNT Proteina morfogenetica ossea/TGFb, regolazione della via di segnalazione	Fattore di rischio moderato
s31	M	56	madre: cancro mammario; padre: cancro gastrico e del colonretto	<i>BRIP1</i> c.139C>G p.Pro47Ala	HRR	Fattore di rischio moderato, eterozigosi per FA
s39	M	71	padre: tumore pancreatico; fratello(s21): polipi	<i>APC</i> c.6588_6589insTAA p.Ser2197 <i>BMPRIA</i> c.1327C>T p.Arg443Cys	Regolazione della segnalazione di WNT Proteina morfogenetica ossea/TGFb, regolazione della via di segnalazione	Poliposi adenomatosa familiare
s42	F	69	madre: leucemia	<i>MLH1</i> c.191A>G p.Asn64Ser	MMR	Fattore di rischio moderato
s69	M	65	padre: cancro alla prostate e ai polmoni	<i>RAD51C</i> c.93delG p.Phe32fs	HRR	Fattore di rischio moderato, eterozigosi per FA

s81	F	71	padre: cancro ai polmoni	<i>BLM</i> c.2098C>T p.Gln700*	HRR	Fattore di rischio moderato
s120	F	58	padre: cancro vescicale o renale	<i>RAD50</i> c.2976_2977delAC p.His993fs	HRR	Fattore di rischio moderato
s140	M	31	nonno paterno: tumore non noto; zia paterna: cancro mammario	<i>CDKN2A</i> c.58del p.Val20*	Regolazione del ciclo cellulare, apoptosi	Sindrome da melanoma familiare con nevi multipli atipici
s141	F	56	madre: cancro intestinale e cancro uterino; padre: cancro intestinale	<i>TSC2</i> c.3856C>T p.Gln1286*	Cell growth regulation	Fattore di rischio moderato
s164	F	62	Non disponibile	<i>MSH6</i> c.2653A>T p.Lys885*	MMR	Sindrome di Lynch
s176	M	71	sorella: cancro al seno e polipi del colon; fratello: polipi del colon	<i>GALNT12</i> c.907G>A p.Asp303Asn	O-glicosilazione	Fattore di rischio moderato
s182	M	80	probando: cancro alla prostata; padre: cancro ai polmoni	<i>FANCG</i> c.1183_1192del p.Glu395fs	HRR	Fattore di rischio moderato, eterozigosi per FA
s185	M	62	padre: cancro a prostate e reni; nonna materna: cancro alla gola	<i>SBDS</i> c.173dup p.Asn59fs	Assemblaggio dei cromosomi	Fattore di rischio moderato
s194	M	65	Sorella: cancro mammario	<i>VHL</i> c.241C>T p.Pro81Ser	Ciclo cellulare, regolazione dell'espressione del fattore inducibile dall'ipossia, degradazione delle proteine	Sindrome di von Hippel-Lindau
s203	M	59	madre: cancro mammario; padre: cancro alla prostata; zio: possibile cancro ai polmoni	<i>NTHL1</i> c.244C>T p.Gln82*  <i>BRCAl</i> c.181T>G p.Cys61Gly	BER HRR	Cancro ereditario del seno e ovarico, eterozigosi per NAP
s204	F	59	padre: cancro a tiroide, polmoni e colon; zio materno: cancro ai polmoni	<i>BRCAl</i> c.5030_5033del p.Thr1677fs	HRR	Cancro ereditario del seno e ovarico
s213	M	71	probando: tumore alla vescica, polmone e retto; fratello: cancro ematologico; sorella: cancro ai polmoni; sorella: cancro al colon	<i>MUTYH</i> c.1103G>A p.Gly368Asp	BER	Fattore di rischio moderato, eterozigosi per MAP
*s242	F	69	Non disponibile	<i>MUTYH</i> c.1103G>A p.Gly368Asp	BER	Fattore di rischio moderato, eterozigosi per MAP
*s243	M	68	Non disponibile	<i>FANCE</i> c.1111C>T p.Arg371Trp	HRR	Fattore di rischio moderato, eterozigosi per FA
s251	M	69	padre: cancro alla protata, nonno materno: possibile cancro al colon	<i>MUTYH</i> c.1103G>A p.Gly368Asp	BER	Fattore di rischio moderato, eterozigosi per MAP
s269	F	48	nonno materno: cancro al colon	<i>FANCM</i> c.3286_3287insT p.Pro1096fs  <i>MUTYH</i> c.1103G>A p.Gly368Asp	HRR BER	Fattore di rischio moderato, eterozigosi per FAe MAP
*s293	M	58	Non disponibile	<i>MSH6</i> c.2570A>C p.Asp857Ala	MMR	Fattore di rischio moderato
*s297	F	53	madre: polipi del colon; padre: sindrome mielodisplastica; nonno materno: cancro ai polmoni	<i>NTHL1</i> c.244C>T p.Gln82*	BER	Fattore di rischio moderato, eterozigosi per NAP

Tabella 22: Varianti identificate nei pazienti con predisposizione genetica. \*: pazienti analizzati durante il tirocinio. Abbreviazioni: BER: riparazione per escissione di basi; HRR: riparazione per ricombinazione omologa; MMR: mismatch repair, FA: Anemia di Fanconi; MAP: poliposi MUTYH-associata; NAP: poliposi NTHL1-associata.

## 6.4 Classificazione dei pazienti e analisi del microbiota

Sono state effettuate due analisi per confrontare il microbiota associato alla mucosa (MAM):

- pazienti mutati (pazienti portatori di varianti patogeniche germinali) vs pazienti sporadici (pazienti senza varianti patogeniche germinali e senza familiarità per CRC);
- pazienti mutati e familiari (pazienti con predisposizione al CRC) vs pazienti sporadici.

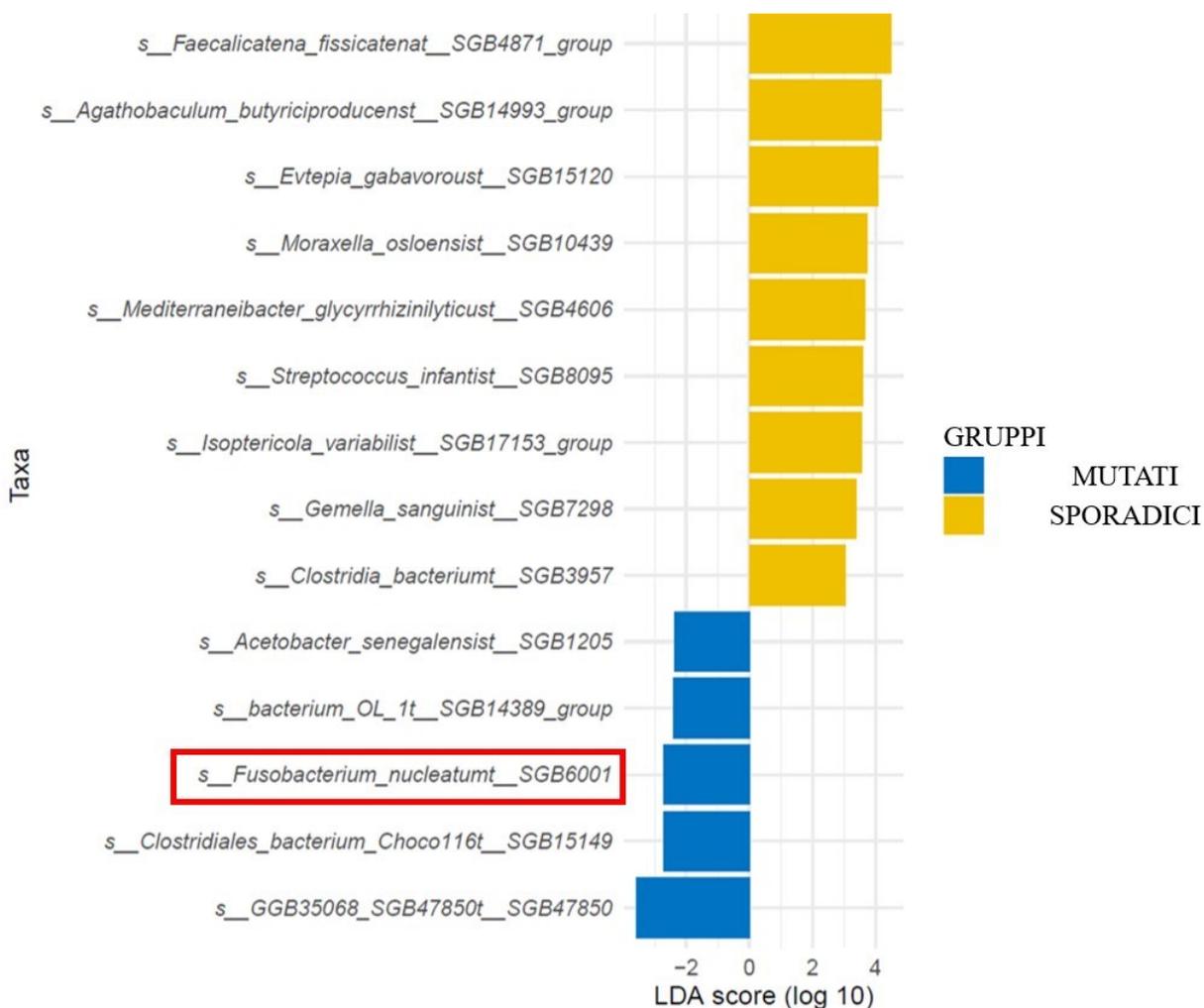
Dei 242 pazienti ne sono stati esclusi 88 per cui il MAM non era disponibile, per un totale di 154 pazienti. La tabella 23 riporta le caratteristiche cliniche dei pazienti sporadici, familiari, mutati e familiari+mutati.

Tabella 23: Caratteristiche cliniche dei pazienti sporadici, familiari, mutati e familiari+mutati.

Caratteristiche cliniche	Pazienti sporadici (N = 63)	Pazienti familiari (N = 71)	p-value	Pazienti mutati (N = 20)	p-value	Pazienti mutati+familiari (N = 91)	p-value
<b>Età in anni</b> Media ( $\pm$ SD)	63.9 ( $\pm$ 7.15)	60.6 ( $\pm$ 10.58)	0.21	63.4 ( $\pm$ 9.66)	0.65	61.2 ( $\pm$ 10.39)	0.37
<b>Istologia</b>							
Adenocarcinoma	3 (4.8%)	9 (12.7%)	0.47	3 (15%)	0.58	12 (13.2%)	0.49
Polipi con alto grado di displasia	26 (41.3%)	25 (35.2%)		8 (40%)		33 (36.3%)	
Polipi con basso grado di displasia	23 (36.5%)	26 (36.6%)		5 (25%)		31 (34%)	
Polipi iperplastici/serrati	6 (9.5%)	8 (11.3%)		2 (10%)		10 (11%)	
Altri	5 (7.9%)	3 (4.2%)		2 (10%)		5 (5.5%)	
<b>Indice di massa corporea</b>							
Normopeso	31 (49.2%)	39 (55%)	0.57	13 (65%)	0.24	52 (57.1%)	0.38
Sovrappeso/obesi	31 (49.2%)	32 (45%)		7 (35%)		39 (42.9%)	
Non disponibili	1 (1.6%)	0		0		0	

Tramite il sequenziamento Shotgun del MAM sono state identificate differenti firme batteriche nei pazienti sporadici e mutati. In particolare, nei pazienti mutati è stato riscontrato un arricchimento di *Fusobacterium nucleatum*, una specie batterica associata al CRC (Ou *et al.*, 2022) (Figura 10).

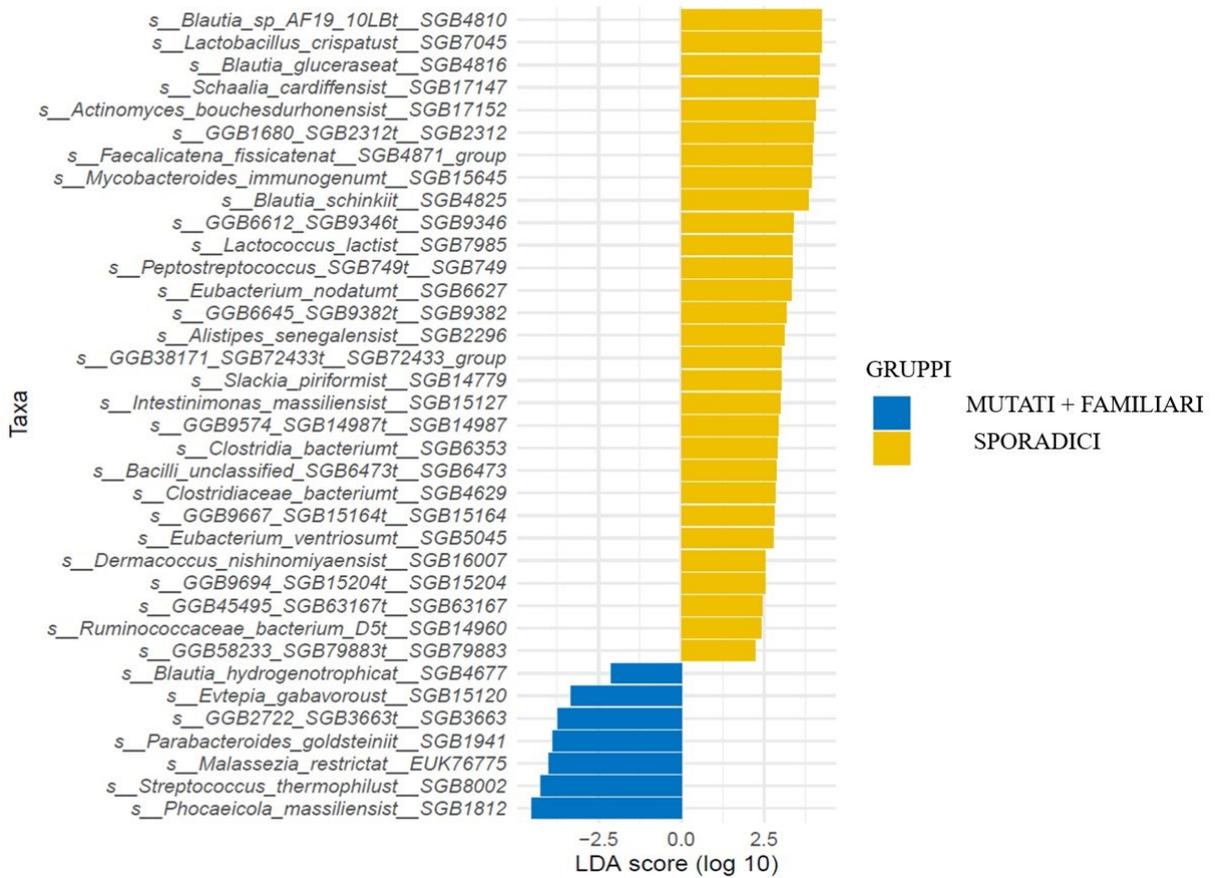
Figura 10: Linear discriminant analysis effect size (LefSe) relativo alle specie batteriche arricchite nel MAM dei pazienti sporadici (giallo) e mutati (blu).



*F. nucleatum* generalmente è associato agli stadi avanzati della cancerogenesi coloretale (Flanagan *et al.*, 2014). La percentuale di pazienti con adenocarcinoma è superiore nei pazienti mutati e familiari rispetto agli sporadici. Abbiamo effettuato la seconda analisi del MAM confrontando i pazienti sporadici con i pazienti mutati e familiari e *F. nucleatum* non è risultato arricchito nel gruppo familiari e mutati vs sporadici. In pazienti con familiarità potrebbero essere stati esposti a fattori di rischio ambientale. Questo suggerisce che l'arricchimento di *F.*

*nucleatum* non sia dovuto agli stadi avanzati della cancerogenesi, ma alla predisposizione genetica del gruppo dei pazienti mutati (Figura 11).

Figura 11: Linear discriminant analysis effect size (LefSe) relativo alle specie batteriche arricchite nel MAM dei pazienti sporadici (giallo) e familiari+mutati (blu).



## **7. Discussione dei risultati**

L'obiettivo di questa tesi è stato studiare l'interazione tra il microbiota intestinale, la predisposizione ereditaria al cancro e la cancerogenesi del colon. In particolare, si è voluto valutare se i pazienti con fattori di rischio ereditari mostrassero un microbiota differente rispetto a pazienti sporadici. Inizialmente, è stato caratterizzato il background genetico di 242 pazienti con polipi del colon-retto e, successivamente, è stato analizzato il microbiota intestinale dei pazienti disponibili. Dalle analisi è emerso che il 10,33% dei pazienti (n=25/242) presentano varianti germinali patogeniche che conferiscono un rischio alto o moderato di sviluppo di CRC. Questa percentuale risulta più alta rispetto a quella attesa in termini di predisposizione genetica al CRC (2-8%) (Valle *et al.*, 2019). Nel nostro studio abbiamo utilizzato lo stesso approccio di Yurgelun e colleghi, che identificarono mutazioni germinali con penetranza moderata o alta nel 9,9 % dei loro pazienti (Yurgelun *et al.*, 2017). Al contrario, altri studi si sono focalizzati solo su mutazioni ad alta penetranza in pazienti ad alto rischio con CRC, cioè pazienti con esordio precoce o storia familiare positiva per CRC o instabilità dei microsatelliti (Yurgelun *et al.*, 2017).

Le implicazioni cliniche del test genetico per tutti i pazienti con CRC sono importantissime in vista delle terapie basate sulla medicina di precisione. A tal proposito, la Food and Drug Administration ha approvato gli inibitori del checkpoint immunitario per il trattamento dei tumori causati dall'alta instabilità dei microsatelliti (MSI-H) o dal deficit del mismatch repair (dMMR) e come trattamento di prima linea di pazienti con CRC metastatico associato a MSI-H o dMMR (Choucair *et al.*, 2021). Inoltre, è noto che individui con cancro alla mammella e cancro ovarico, ma anche al pancreas e alla prostata, che sono portatori di varianti patogeniche ereditarie in *BRCA1* e *2*, mostrano una risposta migliore ai PARP-inibitori, rispetto ai pazienti che non presentano varianti patogeniche germinali (Kaufman *et al.*, 2015). Analogamente, pazienti con CRC che presentano nella linea germinale varianti patogeniche di *BRCA1/2* e altri geni HRR potrebbero trarre benefici da farmaci PARP-inibitori (Yurgelun *et al.*, 2017).

È molto studiato anche il potenziale ruolo terapeutico del microbiota intestinale: l'utilizzo di probiotici (microrganismi vivi con effetto benefico sulla salute dell'ospite) potrebbe ristabilire un microbiota intestinale sano, migliorando la sensibilità ai chemioterapici o ai ICIs (Mori & Pasca, 2021). Per questa ragione, le analisi del microbiota intestinale e la comprensione dei suoi meccanismi e delle sue interazioni risultano sempre più importanti.

Gli studi delle alterazioni del microbiota in pazienti con CRC sporadico o ereditario sono riportati nella rassegna di Mori e Pasca (Mori & Pasca, 2021). Finora, gli studi sul microbiota nelle sindromi di CRC ereditario si sono focalizzati soprattutto su campioni fecali e solo su pazienti con LS, senza confronto con individui non mutati.

Questo lavoro ha l'obiettivo di colmare questa lacuna, mettendo a confronto il microbiota adeso alla mucosa tumorale di pazienti mutati e non mutati.

Sono state identificate differenti firme batteriche tra pazienti mutati e sporadici. Aggiungendo i casi familiari al gruppo dei mutati la composizione batterica risulta differente. Ciò suggerisce che i pazienti mutati sono diversi rispetto a quelli con generica familiarità.

In particolare, è stato riscontrato un arricchimento di *Fusobacterium nucleatum* in pazienti mutati rispetto a quelli sporadici, valutati a livello delle mucose. In letteratura, questo batterio è stato trovato arricchito su biopsie di tessuto tumorale confrontato alla mucosa sana dei pazienti con CRC, e nel microbiota fecale dei pazienti con CRC confrontato ai controlli sani (White & Sears, 2023). Inoltre, ha un ruolo nella cancerogenesi coloretale attraverso diversi meccanismi. Promuove la progressione del tumore attivando la via di segnalazione Wnt/ $\beta$ -catenina tramite l'adesina FadA e per far ciò necessita dell'espressione dell'annessina A1, espressa selettivamente dalle cellule cancerose del colon (Rubinstein *et al.*, 2013; Li *et al.*, 2020).

Rubinstein e colleghi hanno proposto un modello a due colpi per la cancerogenesi coloretale: il primo colpo è rappresentato dall'accumulo di mutazioni somatiche driver nell'ospite e il secondo colpo dall'attività microbica con cui si favorisce la progressione del cancro (Rubinstein *et al.*, 2019). Durante lo studio sono stati trovati diversi pazienti portatori di mutazioni germinali nei geni coinvolti nel controllo della proliferazione cellulare; in questi pazienti la predisposizione germinale e la maggiore abbondanza di *F. nucleatum* sui polipi possono cooperare per favorire la proliferazione cellulare.

Inoltre, *F. nucleatum* induce il danno sul DNA tramite l'adesina FadA (Guo *et al.*, 2020; Karpiński *et al.*, 2022). È stato ipotizzato che *F. nucleatum* induca il danno del DNA nel CRC con l'attivazione FadA-dipendente della via di E-caderina/ $\beta$ -catenina, che porta ad una up-regolazione di *chk2* in topi *ApcMin*<sup>+</sup>. Tuttavia, non è stata stabilita alcuna relazione causale tra FadA e *chk2* (Karpiński *et al.*, 2022). In un altro studio sul carcinoma a cellule squamose della testa e del collo (HNSCC), gli autori hanno dimostrato che *F. nucleatum* promuove la proliferazione di HNSCC sopprimendo l'espressione di *MLH1*, *MSH2* e *MSH6* attraverso l'up-regolazione del microRNA mir-205-5p (Hsueh *et al.*, 2022).

È importante notare che la maggior parte dei geni che abbiamo trovato mutati sono coinvolti nelle vie di riparazione del DNA (HRR, BER, MMR). È stato dimostrato che l'abbondanza di *F. nucleatum* in CRC con instabilità dei microsatelliti (compresi i pazienti con LS e CRC sporadico con MSI) era superiore a quella in CRC sporadici con microsatelliti stabili (MSS) (Ono *et al.*, 2022). Pertanto, *F. nucleatum* e predisposizione genetica possono cooperare per favorire la mutagenesi.

Un altro fattore di virulenza di *F. nucleatum* è Fap2, una lectina che si lega alla D-galattosio-(1-3)-N-acetil-D-galattosamina (Gal-GalNAc), la cui espressione è aumentata nel CRC (Abed *et al.*, 2016). GalNAc e GalGalNAc sono O-GalNAc glicani (Alon-Maimon *et al.*, 2022), componenti importanti delle mucine che compongono la barriera fra il lume intestinale e l'epitelio (González-Morelo *et al.*, 2020).

L'importanza delle anomalie dello strato mucoso nella cancerogenesi del colon è sottolineata dal fatto che il deficit della O-glicosilazione comporta la compromissione della funzione di barriera del muco e favorisce il CRC associato alla colite nei topi, in cui il microbiota promuove l'infiammazione e il cancro (Bergstrom *et al.*, 2016). *GALNT12*, gene che codifica per un enzima che catalizza lo step iniziale della O-glicosilazione, è stato identificato come mutato in 2 pazienti di questo studio. La mutazione di questo gene, che avviene più frequentemente nelle prime fasi della cancerogenesi, determina la riduzione dell'attività enzimatica (37%) (Guda *et al.*, 2009; Pinho *et al.*, 2015). La ridotta attività di *GALNT12* è responsabile della glicosilazione anormale della mucina e aumenta il rischio di CRC (Guda K *et al.*, 2009).

Inoltre, la proteina Fap2 di *F. nucleatum* può legarsi al recettore inibitorio umano TIGIT, riducendo la soppressione delle cellule tumorali da parte delle cellule NK e sopprimendo l'immunità (Karpiński *et al.*, 2022).

Nel complesso, i dati di questo studio supportano l'ipotesi che la tumorigenesi in pazienti con background genetico diverso è favorita da diversi fattori di rischio batterici.

Come estensione del modello a due colpi proposto da Rubinstein e colleghi (Rubinstein *et al.*, 2019), si ipotizza che i batteri possano cooperare con il difetto genetico per promuovere il cancro negli ospiti predisposti a causa di mutazioni germinale. Ad esempio, quando la  $\beta$ -catenina è iperattivata, come nelle cellule con varianti patologiche *APC*, la presenza di *F. nucleatum* può incrementare ulteriormente la sua attività e di conseguenza la proliferazione cellulare.

L'ipotesi di questa interazione nella comunità scientifica è sostenuta da uno studio del Memorial Sloan Kettering Cancer, che sta analizzando il microbiota fecale e la dieta in pazienti con LS o sindromi ereditarie poliposiche del colon, al fine di valutare il ruolo dei batteri intestinali nel rischio di cancro coloretale oltre al rischio ereditario (Clinicaltrials.gov ID: NCT02371135).

In conclusione, è stato identificato un sottogruppo di pazienti con polipi del colon-retto che sono portatori di varianti patogeniche germinali in geni che predispongono al cancro. Inoltre, analizzando il microbiota adeso al polipo, sono state trovate firme batteriche differenti in pazienti mutati rispetto ai pazienti sporadici.

Questi risultati sottolineano l'importanza di caratterizzare la diversa influenza che il microbiota intestinale può avere nella cancerogenesi coloretale sulla base del background genetico di ogni individuo.

## **8. Bibliografia**

- Abed J., Emgård J.E., Zamir G., et al. “Fap2 Mediates *Fusobacterium nucleatum* Colorectal Adenocarcinoma Enrichment by Binding to Tumor-Expressed Gal-GalNAc”. 2016. *Cell Host Microbe* 20, 215-225.
- Abid Z., Cross A.J., and Sinha R. “Meat, dairy, and cancer”. 2014. *The American journal of clinical nutrition* 100 (suppl):386S–93S.
- Aelvoet A.S., Buttitta F., Ricciardiello L., et al.” Management of familial adenomatous polyposis and MUTYH-associated polyposis; new insights”.2022. *Best Practice & Research Clinical Gastroenterology*, 58-59, 101793.
- Alhinai E.A., Walton G.E., and Commane D.M. “The Role of the Gut Microbiota in Colorectal Cancer Causation”. 2019. *Int J Mol Sci* 20, 5295.
- Ali, M. et al. “Characterization of mutant MUTYH proteins associated with familial colorectal cancer”. 2008. *Gastroenterology* 135, 499–507.
- Alon-Maimon, T., Mandelboim, O. & Bachrach, G. “*Fusobacterium nucleatum* and cancer”. 2022. *Periodontol.* 2000 89, 166–180.
- Alteri R., Anderson J., Barnes C., et al. “Colorectal cancer facts and figures 2020-2022”. American Cancer Society.
- Araghi M., Soerjomataram I., Jenkins M., et al. “Global trends in colorectal cancer mortality: projections to the year 2035”. 2019. *Int J Cancer* 144, 2992-3000.
- Baena, R. and Salinas, P. “Diet and colorectal cancer”.2015. *Maturitas* 80, 258–264
- Bailie L., Loughrey M.B. and Coleman E.G. “Lifestyle Risk Factors for Serrated Colorectal Polyps: A Systematic Review and Meta-analysis”. 2017. *Gastroenterology* 152:92–104.
- Barker N. “Adult intestinal stem cells: Critical drivers of epithelial homeostasis and regeneration”. 2014. *Nature Reviews Molecular Cell Biology* 15, 19–33.
- Belcheva A., Irrazabal T., Susan J.R., et al. “Gut microbial metabolism drives transformation of msh2-deficient colon epithelial cells”. 2014. *Cell* 158, 288-299.
- Bergstrom, K. et al. “Defective Intestinal Mucin-Type O-Glycosylation Causes Spontaneous Colitis-Associated Cancer in Mice”. 2016. *Gastroenterology* 151, 152-164.e11.
- Bisgaard M.L., Fenger K., Bülow S., et al. “Familial adenomatous polyposis (FAP): Frequency, penetrance, and mutation rate”. 1994. *Human Mutation* 3, 121–125.

- Boleij A., Hechenbleikner E.M., Goodwin A.C., et al. “The *Bacteroides fragilis* toxin gene is prevalent in the colon mucosa of colorectal cancer patients”. 2015. *Clin Infect Dis* 60, 208-215.
- Bonnet M., Buc E., Sauvanet P., et al. “Colonization of the human gut by *E. coli* and colorectal cancer risk”. 2014. *Clin Cancer Res* 20, 859-867.
- Brenner H. and Chen C. “The colorectal cancer epidemic: challenges and opportunities for primary, secondary and tertiary prevention” 2018. *British Journal of Cancer*, 119:785–792.
- Brinkmeyer M.K. and David S.S. “Distinct functional consequences of MUTYH variants associated with colorectal cancer: Damaged DNA affinity, glycosylase activity and interaction with PCNA and Hus1”. 2015. *DNA Repair* 34, 39–51.
- Burkitt D.P. “Possible Relationships between Bowel Cancer and Dietary Habits”. 1971. *Journal of the Royal Society of Medicine* 64, 964–965.
- Calva-Cerqueira D., Chinnathambi S., Pechman B., et al. “The rate of germline mutations and large deletions of SMAD4 and BMPR1A in juvenile polyposis”. 2009. *Clin Genet* 75, 79-85.
- Casper M., Spier I., Holz R., et al. “Phenotypic variability of MUTYH-Associated polyposis in monozygotic twins and endoscopic resection of a giant polyp in pregnancy”. 2018. *Am. J Gastroenterol* 113, 625-627.
- Castellarin M., Warren R.L., Freeman J.D., et al. “*Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma”. 2012. *Genome Res* 22, 299-306.
- Chaudhry M.S., Gilmour K.C., Housec I.G., et al. “Missense mutations in the perforin (PRF1) gene as a cause of hereditary cancer predisposition”. 2016. *Oncoimmunology* 2;5(7):e1179415.
- Chen C.C., Lin W.C., Kong M.S., et al. “Oral inoculation of probiotics *Lactobacillus acidophilus* NCFM suppresses tumour growth both in segmental orthotopic colon cancer and extra-intestinal tissue”. 2012. *British Journal of Nutrition* 107, 1623-1634.
- Cheng Y., Ling Z. and Li L. “The Intestinal Microbiota and Colorectal Cancer”. 2020. *Front Immunol* 30:11:615056.
- Chong E.S. “A potential role of probiotics in colorectal cancer prevention: review of possible mechanisms of action”. 2014. *World Journal of Microbiol and Biotechnology* 30, 351-74.

- Choucair, K., Radford, M., Bansal, A., Park, R. & Saeed, A. “Advances in immune therapies for the treatment of microsatellite instability-high/deficient mismatch repair metastatic colorectal cancer (Review)”. 2021 *Int. J. Oncol.* 59, 74.
- Copenhagen-Glazer S., Sol A., Abed J., et al. “Fap2 of *Fusobacterium nucleatum* is a galactose-inhibitable adhesin involved in coaggregation, cell adhesion, and preterm birth”. 2015. *Infection and Immunity* 83, 1104-1113.
- Cuevas-Ramos G., Petit C.R., Marcq I., et al. “*Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells”. 2010. *Proc Natl Acad Sci USA* 107, 11537-11542.
- Donohoe D.R., Collins L.B., Wali A., et al. “The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation”. 2012. *Mol Cell* 48, 612-626.
- Esteban-Jurado C., Franch-Expósito S., Muñoz J., et al. “The Fanconi anemia DNA damage repair pathway in the spotlight for germline predisposition to colorectal cancer”. 2016. *Eur J Hum Genet* 24(10):1501-5.
- Fedirko V., Tramacere I., Bagnardi V., et al. “Alcohol drinking and colorectal cancer risk: An overall and dose-Response meta-analysis of published studies”. 2011. *Annals of Oncology* 22, 1958–1972.64.
- Feldmann J., Le Deust F., Ouachèe-Chardin M., et al. “Functional consequences of perforin gene mutations in 22 patients with familial haemophagocytic lymphohistiocytosis “. 2002. *Br J Haematol* 117(4):965-7.
- Flanagan L., Schmid J., Ebert M., et al. “*Fusobacterium nucleatum* associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome”. 2014. *Eur J Clin Microbiol Infect Dis* 33, 1381-1390.
- Friedenreich C. M., Ryder-Burbidge C. and McNeil J. “Physical activity, obesity and sedentary behavior in cancer etiology: epidemiologic evidence and biologic mechanisms”. 2020. *Molecular Oncology* 15(3):790-800.
- Geier M. S., Butler R.N., Howarth G.S. “Probiotics, Prebiotics and Synbiotics: a role in chemoprevention for colorectal cancer?”. 2006. *Cancer Biol Ther* 5(10):1265-9.
- GLOBOCAN 2020: New Global Cancer Data | UICC.
- González-Morelo K.J., Vega-Sagardía M. and Garrido D. “Molecular Insights Into O-Linked Glycan Utilization by Gut Microbes”. 2020. *Front. Microbiol.* 11, 591568.

- Goto, M. et al. “Adenine DNA glycosylase activity of 14 human MutY homolog (MUTYH) variant proteins found in patients with colorectal polyposis and cancer”. 2010. *Hum. Mutat.* 31, E1861-1874.
- Grady W.M., and Carethers J.M. “Genomic and Epigenetic Instability in Colorectal Cancer Pathogenesis”. 2008. *Gastroenterology* 135, 1079–1099.
- Grady W.M., and Pritchard C.C. “Molecular alterations and biomarkers in colorectal cancer”. 2014. *Toxicologic Pathology* 42, 124–139.
- Guda K., Moinoval H., He1 J., et al. “Inactivating germ-line and somatic mutations in polypeptide N-acetylgalactosaminyltransferase 12 in human colon cancers”. 2009. *Proceedings of the national academy of science* 106, 31.
- Guidelines for the Evaluation of Probiotics in Food. FAO/WHO.
- Guo, P. et al. “FadA promotes DNA damage and progression of *Fusobacterium nucleatum*-induced colorectal cancer through up-regulation of *chk2*”. 2020. *J. Exp. Clin. Cancer Res.* CR 39, 2020.
- Gupta, S. et al. “Potential impact of family history-based screening guidelines on the detection of early-onset colorectal cancer”. 2020. *Cancer* 126, 3013–3020.
- Gur C., Ibrahim Y., Isaacson B., et al. “Binding of the Fap2 protein of *Fusobacterium nucleatum* to human inhibitory receptor TIGIT protects tumors from immune cell attack”. 2015. *Immunity* 42, 344-355.
- Han Y.W., Redline R.W., Li M., et al. “*Fusobacterium nucleatum* induces premature and term stillbirths in pregnant mice: implication of oral bacteria in preterm birth”. 2004. *Infect Immun* 72, 2272-2279.
- Han Y.W., Shi W., Huang G.T., et al. “Interactions between periodontal bacteria and human oral epithelial cells: *Fusobacterium nucleatum* adheres to and invades epithelial cells”. 2000. *Infect Immun* 68, 3140-3146.
- Hearle N.C., Rudd M.F., Lim W., et al. “Exonic STK11 deletions are not a rare cause of PeutzJeghers syndrome”. 2006. *Journal of Medical Genetics* 43, e15.
- Housseau F., and Sears C.L. “Enterotoxigenic *Bacteroides fragilis* (ETBF)-mediated colitis in Min (*Apc*<sup>+/-</sup>) mice: a human commensal-based murine model of colon carcinogenesis”. 2010. *Cell Cycle* 9, 3-5.
- Hryhorowicz S., Kaczmarek-Ryś M., Lis-Tanaś e., et al.” Strong Hereditary Predispositions to Colorectal Cancer”. 2022. *Genes* 10;13(12):2326.

- Hsieh C.Y., Osaka T., Moriyama E., et al. “Strengthening of the intestinal epithelial tight junction by *Bifidobacterium bifidum*”. 2015. *Physiol Rep* 3, 12327.
- Hsueh, C.-Y. et al. “*Fusobacterium nucleatum* impairs DNA mismatch repair and stability in patients with squamous cell carcinoma of the head and neck”. 2022. *Cancer* 128, 3170–3184.
- Huang J.Y., Lee S.M., and Mazmanian S.K. “The human commensal *Bacteroides fragilis* binds intestinal mucin”. 2011. *Anaerobe* 17, 137-141.
- Karen Simon. “Colorectal cancer development and advances in screening”. 2016. *Clinical Interventions in Aging*, 19:11:967-76.
- Karpínski, T. M., Ożarowski, M. & Stasiewicz, M. “Carcinogenic microbiota and its role in colorectal cancer development”. 2022. *Semin. Cancer Biol.* 86, 420–430.
- Kaufman, B. et al. “Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation”. 2015. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 33, 244–250.
- Keum N. and Giovannucci E. “Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies”. 2019. *Nat Rev Gastroenterol Hepatol* 16(12):713-732.
- Kostic A.D., Chun E., Robertson L., et al. “*Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment”. 2013. *Cell Host Microbe* 14, 207-15.
- Komine, K. et al. “Functional Complementation Assay for 47 MUTYH Variants in a MutY-Disrupted *Escherichia coli* Strain”. 2015. *Hum. Mutat.* 36, 704–711.
- Kosumi K., Mima K., Baba H., et al. “Dysbiosis of the gut microbiota and colorectal cancer: the key target of molecular pathological epidemiology”. 2018. *J Lab Precis Med* 3, 76.
- Kovacs M.E., Papp J., Szentirmay Z., et al. “Deletions removing the last exon of TACSTD1 constitute a distinct class of mutations predisposing to lynch syndrome”. 2009. *Human Mutation* 30, 197–203.
- Kuwahara A. “Contributions of colonic short-chain Fatty Acid receptors in energy homeostasis”. 2014. *Front Endocrinol* 5, 144.
- Ladabaum U., Dominitz J.A., Kahi C., et al. “Strategies for Colorectal Cancer Screening”. 2020. *Gastroenterology* 158, 418-432.

- Li S.K.H., and Martin A. “Mismatch Repair and Colon Cancer: Mechanisms and Therapies Explored”. 2016. *Trends Mol Med.* 22, 274-289.
- Li, X. et al. “Fusobacterium nucleatum Promotes the Progression of Colorectal Cancer Through Cdk5-Activated Wnt/ $\beta$ -Catenin Signaling”. 2020. *Front. Microbiol.* 11, 545251.
- Ligtenberg M.J.L., Kuiper R.P., Chan, T.L., et al. “Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1”. 2009. *Nature Genetics* 41, 112–117.
- Liu Y., Baba Y., Ishimoto T., et al. “Progress in characterizing the linkage between Fusobacterium nucleatum and gastrointestinal cancer”. 2019. *J Gastroenterol* 54, 33-41.
- Louis P., Hold G.L., and Flint H.J. “The gut microbiota, bacterial metabolites and colorectal cancer”. 2014. *Nat Rev Microbiol.* 12, 661-72.
- Lucas C., Barnich N., and Nguyen H.T.T. “Microbiota, inflammation and colorectal cancer”. 2017. *International Journal of Molecular Sciences* 18, 1310.
- Lung M.S., Trainer A.H., Campbell I., and Lipton, L. “Familial colorectal cancer”. 2015. *Internal Medicine Journal* 45, 482–491.
- Lynch H.T., Lynch P.M., Lanspa S.J., et al. “Review of the Lynch syndrome: History, molecular genetics, screening, differential diagnosis, and medicolegal ramifications”. 2009. *Clinical Genetics* 76, 1–18.
- Lynch H.T., and De la Chapelle A. “Hereditary colorectal cancer”. 2003. *New England Journal of Medicine* 348, 919–932.
- Majumder S., Shivaji U.N., Kasturi R., et al. “Inflammatory bowel disease-related colorectal cancer: “Past, present and future perspectives”. 2022. *World J Gastrointest Oncol.* 14, 547-567.
- Maratt J.K., Stoffel E. “Identification of Lynch Syndrome”. 2022. *Gastrointest Endosc Clin N Am* 32(1):45-58.
- Marchesi J.R., Dutilh B.E., Hall N. et al. “Towards the human colorectal cancer microbiome”. 2011. *PLoS One* 6, e20447 67.
- Mármol I., Sánchez-de-Diego C., Dieste A.P., et al. “Colorectal carcinoma: A general overview and future perspectives in colorectal cancer”. 2017. *International Journal of Molecular Sciences* 18, 197.
- Mauri G., Sartore-Bianchi A., Russo A.G., et al. “Early-onset colorectal cancer in young individuals”. 2018. *Molecular Oncology*, 3(2):109-131.

- McReynolds, L. J. et al. “Risk of cancer in heterozygous relatives of patients with Fanconi anemia”. 2022. *Genet. Med. Off. J. Am. Coll. Med. Genet.* **24**, 245–250.
- Mehta, P. A. & Ebens, C. Fanconi Anemia. in GeneReviews® (eds. Adam, M. P. et al.) (University of Washington, Seattle, Seattle (WA), 1993).
- Molska M., and Reguła J. “Potential Mechanisms of Probiotics Action in the Prevention and Treatment of Colorectal Cancer”. 2019. *Nutrients* **11**, 2453.
- Morgan E., Arnold M., Gini A., et al. “Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN”. 2023. *Gut* **72**,338-344.
- Mori, G. & Pasca, M. R. “Gut Microbial Signatures in Sporadic and Hereditary Colorectal Cancer”. 2021. *Int. J. Mol. Sci.* **22**, 1312.
- Naccarati A., Polakova V., Pardini B., et al. “Mutations and polymorphisms in TP53 gene--an overview on the role in colorectal cancer”. 2012. Oxford University Press (OUP) **27**, 211-8.
- Nielsen M., Hes F.J., Nagengast F.M., et al. “Germline mutations in APC and MUTYH are responsible for the majority of families with attenuated familial adenomatous polyposis”. 2007. *Clinical Genetics* **71**, 427–433.
- Nielsen M., Infante E., Brand R., et al. “MUTYH Polyposis”. 2012. *Gene reviews*.
- Nookala R.K., Hussain S. and Pellegrini L.” Insights into Fanconi Anaemia from the structure of human FANCE”.2007. *Nucleic Acids Res* **35**(5):1638-48.
- Olkinuora A. P., Peltomaki P. T., Aaltonen L. A., et al. “From APC to the genetics of hereditary and familial colon cancer syndromes “. 2021. *Hum. Mol. Genet.* **30**, R206–R224.
- Ono, T. et al. “Fusobacterium nucleatum load in MSI colorectal cancer subtypes”. 2022. *Int. J. Clin. Oncol.* **27**, 1580–1588.
- Ou S., Wang H., Tao Y., Luo K., et al. “Fusobacterium nucleatum and colorectal cancer: From phenomenon to mechanism cellular and infection microbiology” 2022. *Front Cell Infect Microbiol* **12**:1020583.
- Pan P., Yu J., Wang L.S. “Colon Cancer: What We Eat”. 2018. *Surgical Oncology Clinics of North America* **27**, 243–267.
- Parsa, F. G. et al. “Fanconi Anemia Pathway in Colorectal Cancer: A Novel Opportunity for Diagnosis, Prognosis and Therapy”.2022. *J. Pers. Med.* **12**, 396.
- Pinho, S. S. & Reis, C. A. “Glycosylation in cancer: mechanisms and clinical implications”. 2015. *Nat. Rev. Cancer* **15**, 540–555.

- Puccini A., Berger M.D., Naseem M, et al. “Colorectal cancer: epigenetic alterations and their clinical implications”. 2015. *Biochimica et Biophysica Acta - Reviews on Cancer* 1868, 439–448.
- Risma, K. A., Frayer, R. W., Filipovich, A. H. & Sumegi, J. “Aberrant maturation of mutant perforin underlies the clinical diversity of hemophagocytic lymphohistiocytosis.”.2006. *J. Clin. Invest.* 116, 182–192.
- Rizzolo P., Silvestri V., Bucalo A., et al.“Contribution of MUTYH Variants to Male Breast Cancer Risk: Results From a Multicenter Study in Italy “. 2018. *Front Oncol* 4:8:583.
- Rubinstein M.R., Wang X., Liu W., et al. “Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/ $\beta$ -catenin signaling via its FadA adhesin”. 2013. *Cell Host Microbe* 14,195-206.
- Rubinstein, M. R. et al. “Fusobacterium nucleatum promotes colorectal cancer by inducing Wnt/ $\beta$ -catenin modulator Annexin A1”.2019. *EMBO Rep.* 20, e47638.
- Ruggieri, V. et al. “Loss of MUTYH function in human cells leads to accumulation of oxidative damage and genetic instability”. 2013. *Oncogene* 32, 4500–4508.
- Sah B.N., Vasiljevic T., McKechnie S., et al. “Effect of probiotics on antioxidant and antimutagenic activities of crude peptide extract from yogurt”. 2014. *Food Chem* 156, 264-270.
- Samadder N.J., Jaspersen K.N., and Burt R.W. “Hereditary and Common Familial Colorectal Cancer: Evidence for Colorectal Screening”. 2015. *Digestive Diseases and Sciences* 60, 734–747.
- Scott A.J., Alexander J.L., Merrifield C.A., et al. “International Cancer Microbiome Consortium consensus statement on the role of the human microbiome in carcinogenesis”. 2019. *Gut* 68, 1624-1632.
- Shanahan F. “Inflammatory bowel disease: immunodiagnostics, immunotherapeutics, and eotherapeutics”. 2001. *Gastroenterology* 120, 622-635.
- Siege R., Sandeep Wagle N., Cercek A., Robert A. et al. “Colorectal cancer statistics, 2023”. 2023, *A cancer journal for clinicians*, 73(3):233-254.
- Sieni, E. et al. “Familial hemophagocytic lymphohistiocytosis may present during adulthood: clinical and genetic features of a small series”.2012. *PloS One* 7, e44649.
- Song M., Chan A.T., and Sun J. “Influence of the Gut Microbiome, Diet, and Environment on Risk of Colorectal Cancer”. 2020. *Gastroenterology* 158, 322-340.

- Song M., Garrett W.S., and Chan A.T. “Nutrients, foods, and colorectal cancer prevention”. 2015. *Gastroenterology* 148, 1244-1260.
- Srikanth C.V., and McCormick B.A. “Interactions of the intestinal epithelium with the pathogen and the indigenous microbiota: a three-way crosstalk”. 2008. *Interdiscip Perspect Infect Dis* 2008, 626827.
- Stoffel, E.M., and Kastrinos F. “Familial colorectal cancer, beyond lynch syndrome”. 2014. *Clinical Gastroenterology and Hepatology* 12, 1059–1068.
- Sung H., Ferlay J., Siegel R.L., et al. “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries”. 2021. *CA Cancer J Clin* 71, 209-249.
- Syngal, S. et al. “ACG clinical guideline: Genetic testing and management of hereditary gastrointestinal cancer syndromes”. 2015. *Am. J. Gastroenterol.* **110**, 223–262.
- Thanikachalam K., and Khan G. “Colorectal Cancer and Nutrition”.2019. *Nutrients* 14;11(1):164.
- Testa U., Pelosi E., and Castelli G. “Colorectal Cancer: Genetic Abnormalities, Tumor Progression, Tumor Heterogeneity, Clonal Evolution and Tumor-Initiating Cells”. 2018. *Medical Sciences* 6, 31.
- Tian, Y. et al.” Calculating the Starting Age for Screening in Relatives of Patients With Colorectal Cancer Based on Data From Large Nationwide Data Sets”.2020. *Gastroenterology* 159, 159-168.e3.
- Tjalsma H., Boleij A., Marchesi JR., et al. “A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects”. 2012. *Nat Rev Microbiol* 10, 575-582.
- Toyota M., Ahuja N., Ohe-Toyota M., et al. “CpG island methylator phenotype in colorectal cancer”. 1999. *Proceedings of the National Academy of Sciences of the United States of America* 96, 8681–8686.
- Uchida S., Kojima T. and Sugino T. “Frequency and Clinicopathological Characteristics of Patients With KRAS/BRAF Double-Mutant Colorectal Cancer: An In Silico Study”.2022. *Pathology Oncology research*, 28:1610206.
- Valle, L. et al. “Update on genetic predisposition to colorectal cancer and polyposis.” 2019. *Mol. Aspects Med.*69, 10–26.
- Voskoboinik I., Smyth M. J. and Trapani J. A. “Perforin-mediated target-cell death and immune homeostasis”. 2006. *Nat Rev Immunol* (12):940-52.

- Wang T., Cai G., Qiu Y., et al. “Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers”. 2012. *ISME J* 6, 320-9.
- Wei H., Dong L., Wang T., et al. “Structural shifts of gut microbiota as surrogate endpoints for monitoring host health changes induced by carcinogen exposure”. 2010. *FEMS Microbiol Ecol* 73, 577-86.
- Weren R.D., Ligtenberg M.J., Kets C.M., et al. “A germline homozygous mutation in the baseexcision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer”. 2015. *Nat Genet* 47, 668-71.
- White, M. T. & Sears, C. L. “The microbial landscape of colorectal cancer”.2023. *Nat. Rev. Microbiol.* doi:10.1038/s41579-023-00973-4.
- Win, A. K. et al. “Cancer risks for monoallelic MUTYH mutation carriers with a family history of colorectal cancer”.2011. *Int. J. Cancer* 129, 2256–2262.
- Wolin K.Y., Yan Y., Colditz G.A., et al. “Physical activity and colon cancer prevention: A metaanalysis”. 2009. *British Journal of Cancer* 100, 611–616.
- Wong M.C.S., Huang J., Lok V., et al. “Differences in Incidence and Mortality Trends of Colorectal Cancer Worldwide Based on Sex, Age, and Anatomic Location”.2021. *Clinical Gastroenterology and Hepatology*, 19:955–966.
- Wu N., Yang X., Zhang R., et al. “Dysbiosis signature of fecal microbiota in colorectal cancer patients”. 2013. *Microb Ecol* 66, 462-70.
- Xi Y., and Xu P. “Global colorectal cancer burden in 2020 and projections to 2040”. 2021. *Transl Oncol* 14, 101174.
- Yamagishi H., Kuroda H., Imai Y., et al. “Molecular pathogenesis of sporadic colorectal cancers”. 2016. *Chinese Journal of Cancer* 35, 4.
- Yang J., McDowell A., Kim E.K., et al. “Development of a colorectal cancer diagnostic model and dietary risk assessment through gut microbiome analysis”. 2019. *Exp Mol Med* 51, 1–15.
- Yurgelun, M. B. et al. “Cancer Susceptibility Gene Mutations in Individuals With Colorectal Cancer”.2017. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 35, 1086–1095.
- Zeng H., Hamlin SK., Safratowich BD., et al. “Superior inhibitory efficacy of butyrate over propionate and acetate against human colon cancer cell proliferation via cell cycle arrest and apoptosis: linking dietary fiber to cancer prevention”. 2020. *Nutr Res* 83, 63-72.

- Zhao H., Ming T., Tang S., et al. “Wnt signaling in colorectal cancer: pathogenic role and therapeutic target”. 2022. *Mol Cancer* 21, 144.